HABITAT SUITABILITY MAPPING OF THE AMERICAN LOBSTER FOR USE IN MARINE SPATIAL PLANNING

by

Anne McKee

Submitted in partial fulfillment of the requirements
for the degree of Master of Science

at

Dalhousie University
Halifax, Nova Scotia
July 2018

Dedicated to Tom.

Thank you.

TABLE OF CONTENTS

LIST OF TABLES

List of Figures

ABSTRACT

Marine spatial planning (MSP) is a management tool which could help mitigate the conflict that exists between the American lobster fishery and the net-pen salmon aquaculture industry in the Canadian Maritime provinces. However, lobster habitat suitability maps, which are a necessary feature of these MSP, have not been created in most areas. This thesis presents two studies which demonstrate acoustic-based methods of developing habitat suitability maps for the American lobster for use in MSP. The first study demonstrates success with the method for adult lobsters, and highlights the importance of explicitly analysing spatial scale and resolution in benthic habitat models. The second study explores the same acoustic method in tandem with juvenile lobster trawl data, and demonstrates that juveniles live on fine sediments with no preference between the substrate categories. This suggests that a deeper understanding of juvenile lobster habitat is needed to fully map habitat suitability for MSP.

List of Abbreviations Used

Abbreviation	Description
AGDS	Acoustic ground discrimination system
AUC	Area under the curve
BC	Boulder-Cobble
BG	Boulder-Gravel
BS	Bedrock-Sediment
CG	Cobble-Gravel
CL	Carapace length
DGPS	Differential global positioning system
FN	False negative
FP	False positive
GPS	Global positioning system
GS	Gravel-Sand
GV	Gravel
HG	Heterogeneous
LFA	Lobster Fishing Area
LLWLT	Lower Low Water Large Tide
MD	Mud
MP	Mean point (binned pings)
MSP	Marine Spatial Planning
NPV	Negative predictive value
POP	Probability of presence
PPV	Positive predictive value
ROC	Receiver operating characteristics
ROR	Relative occurrence rate
RK	Rock
SA	Sand
SB-AGDS	Single beam acoustic ground discrimination system
SBES	Single beam echosounding system

Abbreviation	Description
SDM	Species distribution model
TN	True negative
TP	True positive
UNBSJ	University of New Brunswick in Saint John

ACKNOWLEDGEMENTS

Firstly, I want to acknowledge that my work was completed on the ancestral and unceded territory of the Mi'kmaq, Wolastoqiyik (Maliseet), and Peskotomuhkati (Passamaquoddy) peoples. As the original owners and custodians of these lands, I wish to thank them and pay respect to the traditional knowledge embedded within the Indigenous stewardship of this country.

Next, I'd like to thank my supervisor, Jon Grant, and my advisory committee, Anna Metaxas and Chris Algar, for their advice and guidance in helping me accomplish my goals. Their support was essential to the completion of this research and I am grateful to have had it. I would also like to thank Ramón Filgueira for his participation as my external examiner and final member of my examination committee. Additional thanks to Chris Taggart for the extra advice and support provided along the way.

Thank you to Francisco Bravo, Danni Harper, Chantal Giroux, and Hart Koepke for the support and companionship through the bulk of my time in the Grant lab, and to Leigh Howarth, Stephen Ennis, Caitlin Stockwell, and Meredith Burke for making the last stretch that much more enjoyable. Thank you to Myriam Lacharité for answering my rookie questions about mapping and spatial analysis. A special thank you goes to Jeff Barrell for all of his invaluable expertise and support through the years – I would have been much more adrift without your help.

My extensive fieldwork would have been impossible if not for the people who supported me through it, and I am forever grateful to them all. Thank you to Rémy Rochette, Kristin Dinning, and the rest of the benthic ecology lab at UNBSJ for their collaborative efforts, their patience with my changing methods, and for letting me tag along on their surveys. Thank you to the multitude of fishermen, boat operators, and Cooke staff who helped me collect my data, particularly Jim Fraelic, Vincent Fisher, Scott Leslie, Jeff Nickerson, and the crew of the Fundy Spray.

Thank you to the Department of Oceanography, both the students and the administrative staff; a particular thank you to Lori Lawton for always knowing who to call, what to do, or who to call to figure out what to do. Thanks to Hansen Johnson and, again, Hart Koepke for being my office-based sounding boards. Thank you to Lisa Arblaster for commiserating over ArcGIS and offering welcome distractions. And a huge thank you to my teammates at Bushido Kai MXT Halifax – aggressive cuddling turns strangers into family.

And finally, to Tom Nightingale: there are no words to properly express how grateful I am that you are in my life. Thank you.

Chapter 1

Introduction

The growth of the aquaculture industry in the Canadian Maritime provinces has led to increased conflict and tension between various groups of interest, particularly the lobster fishery and the net-pen salmon aquaculture industry. Concerns about lobster population health, trapping access, and water pollution divide rural fishing communities' support for aquaculture (*Wiber et al.*, 2012), straining relationships and muting communication between the groups. Marine spatial planning (MSP) is a management tool that could be used to mitigate some of these tensions by organising the use of the coast into a mutually beneficial and efficient structure, creating physical distance between the two industries on the water. However, the crucial layers of lobster habitat data have not been created in many of the locations where they would be most useful.

This "problem of the missing map" is prevalent in many potential MSP data sources, from sociological and human geographical concerns (*St.Martin and Hall-Arber*, 2008), to species growth patterns (*Harris and Stokesbury*, 2006), to ecological population and community models (*Crowder and Norse*, 2008). This is especially true at the finer, local scales used in areas smaller than regional (*Holmes et al.*, 2008). The missing data layers are essential to the development of complete and cohesive marine spatial planning programs and their absence can result in ineffective plans, but creating these map layers is not a trivial task. In particular, species habitat maps can be difficult to make because the concept of habitat is a relatively flexible one that can change with the circumstances of the population being studied.

For species that inhabit the benthic environment for some portion of their life, such as the American lobster, there is the potential to use the surficial geology of the seafloor to build a spatial understanding of the species distribution. Many species use particular substrate types as habitat (Atlantic and Greenland cod (*Laurel et al.*, 2004); barnacles (*Pineda and Caswell*, 1997); hermit crabs (*Bertness*, 1981)); and American lobsters (*Tremblay et al.*, 2009), and with appropriate sampling and data collection, these preferences can be exploited to create habitat suitability maps. However, accurate collection of these data can be extremely difficult. In this thesis I investigated two different species sampling methods and the use of a single beam acoustic ground discrimination system (SB-AGDS), which is a relatively inexpensive and easy-to-use tool designed to report the benthic substrate, and discussed the advantages and disadvantages in using these methods as tools for creating lobster habitat maps for use in future marine spatial planning programs.

1.1 Outline and Objectives

The goal of this thesis was to explore the applicability, limitations, and accuracy of using a SB-AGDS-based approach to creating lobster habitat suitability maps for the future use in MSP. This goal was reached through two branches of investigation, each represented here as a chapter in this thesis.

Chapter 2 focuses on adult lobsters in Liverpool Bay, Nova Scotia, using a combination of acoustic, video, and lobster trap data to create two versions of a habitat map. More specifically, this chapter was designed to explore the importance of spatial analysis on habitat mapping and the effects spatial resolution can have on interpretation.

Chapter 3 addresses juvenile lobsters and the unique limitations they present in predicting their habitat in Maces Bay, New Brunswick. This chapter takes on the specific challenge of modelling their distribution on less preferred substrates, and examines the implications presented by the models.

Finally, the two studies are summarised in a conclusion chapter that includes recommendations on modifying, explanding and applying the methods demonstrated in the document in other regions.

CHAPTER 2

DEVELOPING HABITAT MAPS OF ADULT AMERICAN LOBSTERS FOR USE IN MARINE SPATIAL PLANNING

2.1 Introduction

Coastal marine waters in Atlantic Canada are used for a variety of activities, including fishing, aquaculture, resource extraction, transport, shipping, and recreation. The close proximity and occasional overlap of some activities in these coastal waters can lead to conflict between users (*Wiber et al.*, 2012; *Marshall*, 2001; *Ivany et al.*, 2014). Marine spatial planning (MSP) is a planning framework that is designed to manage such user-user conflicts through structuring the spatial and temporal distribution of ocean-based activities in an efficient manner that is sensitive to the needs of ecosystem, economic, and social objectives (*Foley et al.*, 2010). However, one of the major flaws with MSP is a chronic lack of map layers that detail the required data in suitable scales (*St.Martin and Hall-Arber*, 2008; *Holmes et al.*, 2008; *Crowder and Norse*, 2008; *Harris and Stokesbury*, 2006). Specifically of interest to this study, spatial habitat representations for species of note are important components that are frequently missing from the MSP process.

Commercial fisheries and net-pen aquaculture frequently share coastal space in Atlantic Canada. In Nova Scotia, the American lobster (*Homarus americanus*) is the most valuable commercial catch, with landings totalling to almost $730 million in 2016 (*DFO*, 2016). However, in-shore lobster fishing grounds are often used as farming sites for Atlantic salmon (*Salmo salar*). This overlap can cause conflict in local fishing communities, particularly regarding concerns about water pollution, lobster stock health, and trapping access (*Wiber et al.*, 2012; *Marshall*, 2001; *Ivany et al.*, 2014). MSP strategies can be applied here through the delineation of local benthic lobster habitat, thereby allowing the placement of aquaculture sites in a manner that minimizes their overlap with lobster habitat. Describing the lobster habitat can be accomplished through species distribution models (SDMs). SDMs, which are also referred to as habitat suitability models, are a common type of habitat map which describe the suitability of an environment to support a species by representing where mapped environmental conditions, such as substrate, match the niche conditions of the species (*Brown et al.*, 2011; *Franklin*, 2010).

Adult American lobsters consistently demonstrate preferential selection for shelter-forming substrates in the benthic marine environment. The most strongly preferred substrate is complex, pre-formed shelter such as boulders and cobble (*Tremblay et al.*, 2009; *Selgrath et al.*, 2007; *Bologna*, 1993; *Spanier*, 1993), but shelter can also be constructed as burrows in stable mud (*Spanier*, 1993; *Lawton and Lavalli*, 1995; *Cooper and Uzmann*, 1980) or shallow bowl-shaped depressions in sand (*Cooper and Uzmann*, 1980). The availability of boulders has an effect on the abundance of adult lobsters living on less preferred substrate; they will live on mud in areas where boulders are scarce, on sand if both mud and boulders are scarce (*Cooper and Uzmann*, 1980; *Spanier*, 1993; *Bologna*, 1993), and will move into artificial reefs that are introduced on mud or sand barrens (*Spanier*, 1993; *Bologna*, 1993). However, detection and mapping of these potential habitat characteristics can be difficult for researchers due to the water that covers the seafloor. A common approach to resolving this issue is through echosounding, which is a form of acoustic remote sensing that detects the physical attributes of the seafloor substrate through high frequency sound pulses. Both multibeam (MBES) and single beam (SBES) echosounder systems have been used in tandem with acoustic ground discrimination systems (AGDS; when used with SBES, SB-AGDS), which group the echosounder signals according to their characteristics. These systems are commonly used to map shelf and coastal seafloor substrate distributions and have been used for the purpose of providing environmental layers to SDM for decades (*Brown et al.*, 2011; *Freitas et al.*, 2011). It is this approach which I used in the following study to model habitat for *H. americanus* in a small bay in Atlantic Canada for the purpose of lessening spatial overlap between the lobster fishery and aquaculture via MSP.

2.2 Materials and Methods

2.2.1 Study Site

Liverpool Bay is located approximately 100 km southwest of Halifax, Nova Scotia (44°2' N, -64°40' W), on the Atlantic coast (Figure 2.1). A narrow bay at 4.5 km long and 2.6 km wide at the maximum, it is the exit point of the Mersey River, which extends trumpet-like into the ocean to form the bay. Coffin Island, a small island of approximately 0.75 km^2 that is located 1.5 km northeast of the mouth of the bay, provides some shelter to the Cooke Aquaculture salmon pens tucked between the island and the mainland. The town of Liverpool is at the head of the bay and the village of Brooklyn is on the north shore; the bulk of the activity on the bay is from commercial fishing. The research was focused in the eastern half of the bay, partially due to the location of the fish farm and partially due to the placement of lobster traps. The tidal range of the bay is approximately 1-2 m.

2.2.2 Data Collection

2.2.2.1 Acoustic Data

Acoustic data were collected using a vertically-oriented 204.8 kHz 8.6° beam angle transducer and a BioSonics Inc. MX aquatic habitat echosounder. This single-beam system is designed specifically to detect, classify, and map substrate in coastal areas. The local fishing boat on which the echosounder system was mounted was kept to a speed of 5 kts (2.6 ms^{-1}) to minimize noise interference and to maintain consistent coverage. The rate of acoustic ping production was 5 Hz, creating in a sequential distance between pings of approximately 0.5 m, which were georeferenced with DGPS with a positional accuracy of approximately 2 m. The data were collected over the course of two consecutive days in mid November of 2016. Day 1 included roughly parallel east-west transects separated by approximately 100 m (range 80-120 m). Day 2 had parallel tracks midway between Day 1 transects, resulting in data coverage over the entire survey area at 50 m transect spacing (Figure 2.2). Track separation distances were chosen based on the size of the survey areas and practicalities of completion within a single day. The area covered amounted to approximately 6.3 km^2 and 196,817 data points were collected in total.

2.2.2.2 Ground-Truth Video Data

Ground-truthing data for seafloor substrate were collected via a drop video camera (Seaviewer 950 Series, colour version), and were later used in verification of the acoustic

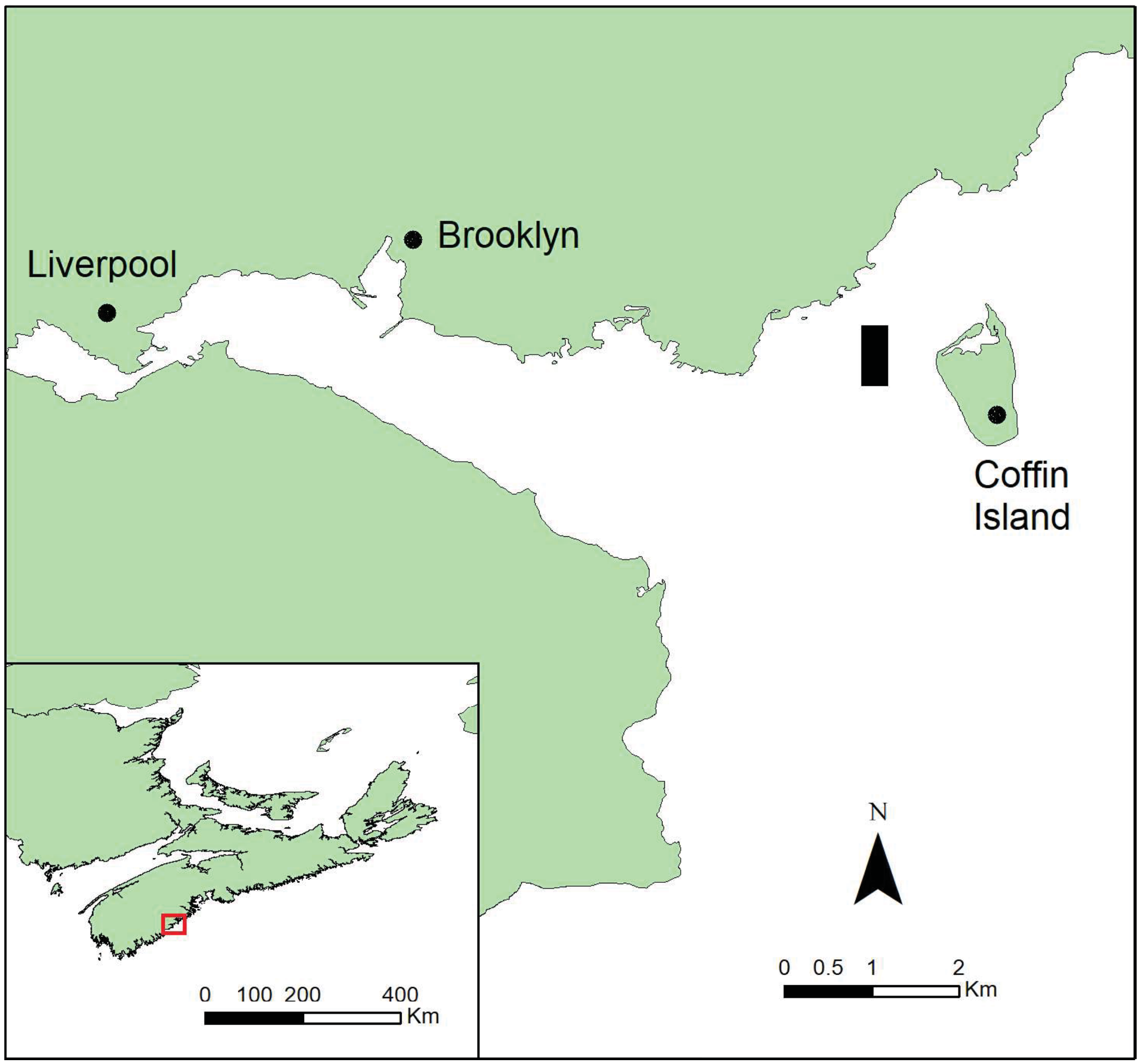

Figure 2.1: Map of Liverpool Bay. (*Inset*) Location within Nova Scotia. The solid black rectangle represents the location of the salmon net pens.

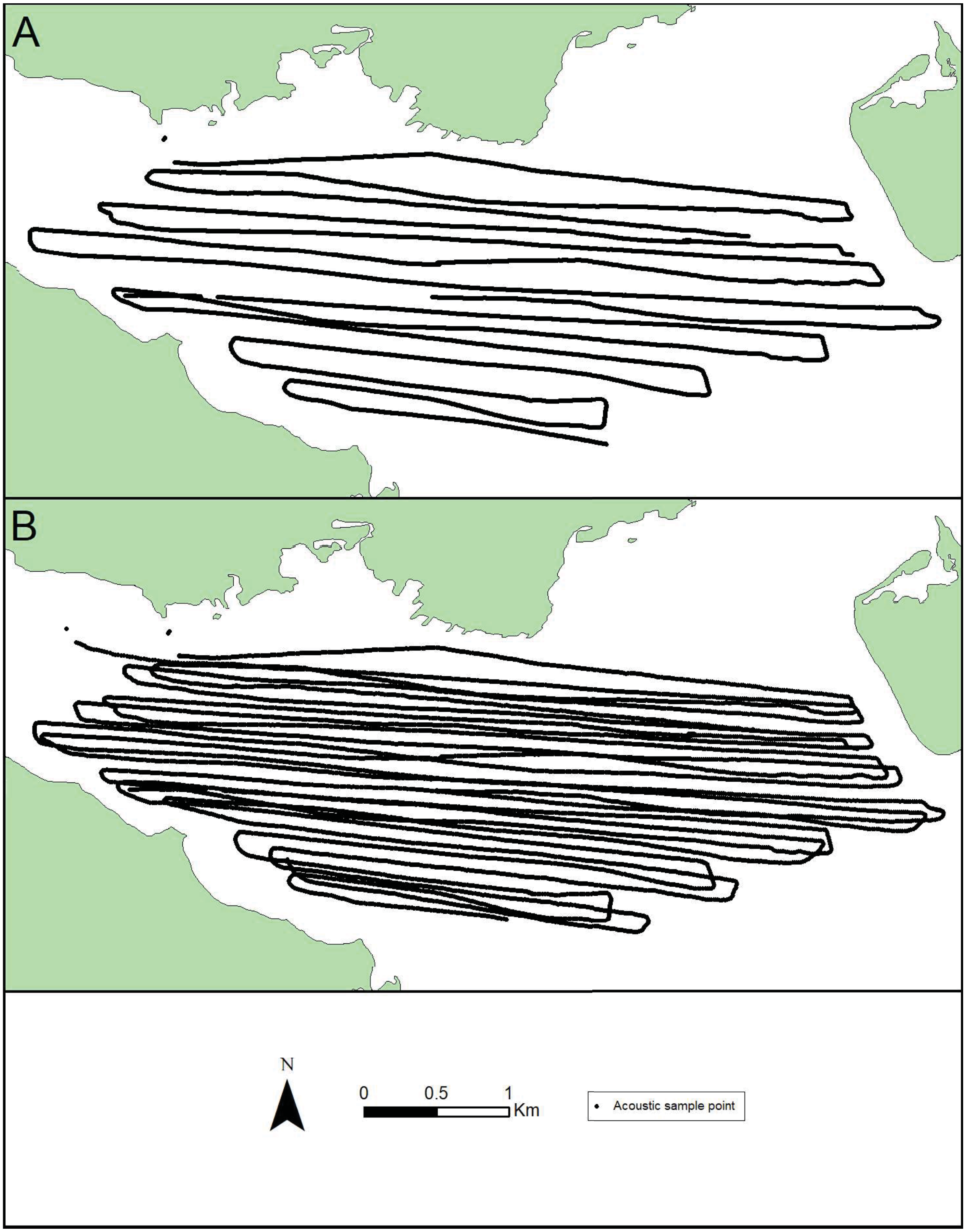

Figure 2.2: The acoustic data survey, with the tracks spaced 100 m apart (A) and 50 m apart (B).

data. Prior additional videos had been taken in the same area in July and September of 2015. Videos were taken both on top of and between the acoustic transects and were georeferenced with a GPS with a positional accuracy of approximately 2-3 m.

2.2.2.3 Trap Data

The lobster fishery in Liverpool Bay is part of Lobster Fishing Area (LFA) 33, where the fishing season is open from the end of November to the end of May. The locations of the lobster trap buoys were used as a surrogate for the presence of lobsters based on fishers' experience in knowing where to deploy. Individual traps or trap strings are marked by distinctive floats and no other similar buoys are present in the bay. The trap data were collected using a DJI Phantom 4 aerial drone during two days in May of 2017, flown within the area covered by the acoustic data collection. The drone took 12.4 megapixel photos at a rate of 0.2 Hz at an altitude of 90 m (2089 photos). This created an overlap between the photos to ensure complete coverage. The drone was configured to always face north (gimbal yaw of 0°), such that the top of the resulting photographs would also correspond with north. The camera was configured to face straight down (gimbal pitch of 90°) to reduce unnecessary complexity in later analysis.

2.2.3 Data Analysis and Substrate Map Creation

2.2.3.1 Acoustic Data

The acoustic data was replicated into six sets to account for the two track separation distances (ie. 100 m and 50 m) and the three spatial resolutions (ie. 2.5 m, 5 m, and 10 m). Resolution was chosen based on the accuracy of the DGPS (~2m horizontal), approximate distance between sequential data points, and variation in the ping footprint with changing depth. To create the varying resolutions, pings were binned at intervals of either 5, 10, or 20, and these bins were referred to as mean points (MPs). The sets were then assigned names according to their combination of track separation and resolution, such that the set with 100 m separation and 5 m resolution was *100-10* (0.5 m between pings, 10 pings binned for 5 m resolution).

Cluster Analysis: For each of the six sets, data were cleaned and trimmed by removing empty pings and pings with <0.25 m sequential spacing, and applying a tidal correction based on the Lower Low Water Low Tide (LLWLT) datum. The data were then analysed in Biosonics VisualHabitat v2.0.1, which conducted an unsupervised fuzzy centroid means cluster analysis to group the MPs into clusters based on substrate characteristics captured at each point. This analysis provided a cluster membership likelihood for each MP, indicating how well each MP fit the description of each cluster. The primary cluster memberships of the MPs were in the clusters where they best fit. The analysis also reported the percentage of the total number of MPs that fit into each cluster, revealing which clusters were rare or likely constructed of error characteristics. This information was then processed in Matlab to calculate the second best fitting cluster for each MP, providing a means by which to compare clusters – if two clusters shared a large number of the same MPs between their primary and secondary memberships, it is likely that the two clusters represented the same substrate and were combined into a single cluster. Since the cluster analysis was conducted to deliberately overestimate the number of clusters in an attempt to avoid accidentally conflating potential substrates, this paring down and combining of the clusters was an important step and it was completed in tandem with visual inspection and interpretation of the echograms. Between three and five final clusters were finalised for each of the six sets of data.

Substrate Categories: The assignment of substrate categories to clusters was based on four substrate categories developed for lobster habitat in SW Nova Scotia using the Wentworth scale (*Tremblay et al.*, 2009)(Table 2.1). These categories and the method of creating them proved useful for the purposes of this thesis work, but a number of changes were made in order to tailor the system to better suit the requirements of an acoustic data set. A series of early analysis attempts indicated that the acoustic method was not sensitive enough to differentiate between BC and CG substrates, and was unable to consistently and accurately cluster the MPs appropriately, so BC and CG were combined into a single category, Rock (RK). Similarly, BG was rarely discovered through the acoustic data due to its sporadic nature, so it was removed as a category. The GS category was split into three categories: Gravel (GV), Sand (SA), and Mud (MD). This expansion was done in part because adult lobsters display preferences between those categories and in part because the separate acoustic signals of Gravel, Sand, and Mud are distinct enough to differentiate – GV was similar to RK but smoother, SA was smooth and hard (between -40 dB and -25 dB), and MD was smooth and soft (-30 dB to -10 dB). The differences between GV, SA, and MD are also visible in the underwater video, such that GV was categorised according to the Wentworth scale (Table 2.1), and SA and MD were differentiated based on the presence and absence of sedimentary ripples (*Whitlatch*, 1977), respectively. This allowed for confirmation of the acoustic data. Mixed substrate, where more than one category was present within the scope of the video, was classified as the largest substrate type present. The final substrate categories used in the further analysis were RK, GV, SA, and MD.

Table 2.1: A) The grain sizes of the sediments and B) the substrate categories used by *Tremblay et al.* (2009).

A.

Sediment	Grain Size Range
Sand	<0.4 cm
Gravel	0.4 – 6 cm
Cobble	6 – 26 cm
Boulder	>26 cm

B.

Acronym	Sediments
BC	Boulder, Cobble
BG	Boulder, Gravel
CG	Cobble, Gravel
GS	Gravel, Sand

2.2.3.2 Ground-Truth Video Data

These four substrate categories were then used to categorise the substrate visible on the underwater video, which was sectioned into 132 individual video segments by GPS location. Each of these segments was visually examined and assigned the substrate category that best applied. To assign substrate categories to the clusters, a subsample of approximately 20% of the video data (25 of 132 segments) was randomly selected from the videos taken on top of the acoustic transects, so that the segments shared GPS locations with acoustic data. This allowed the comparison of the substrate-assigned segments to the cluster-assigned MPs, connecting clusters with substrate categories through the shared location data. The previous analysis resulted in one of the data sets having five clusters (versus the other five sets having only three), but the comparison of the MPs with the video segments showed overlap in the clusters that had not previously been revealed and reduced the number to three. The mismatch between three clusters and four substrate categories was solved by the removal of MD as a category; only three of the video segments were assigned MD, which indicates that it is not a common substrate and likely would not have shown up in the cluster analysis.

2.2.3.3 Substrate Maps

Indicator kriging was used to predict the distribution of substrate categories within the surveyed area by quantifying the spatial autocorrelation of the acoustic data. The semi-variograms were defined using a spherical model, automatic detection of parameters, and a smoothing factor of 0.2. There were no identified anisotropies. The kriging produced raster map layers that displayed the probability of the grid cells belonging to each of the three substrate categories, and a single layer was created where the cells were assigned the substrate categories that had the highest probability. This process was repeated for all six data sets using ArcMap v10.5 in the WSG84 datum.

A depth raster of the survey area was created through the interpolation of the depth data retrieved from the acoustic data. It was converted into 1 m depth contours, which were applied to the substrate maps.

2.2.4 Validation and Uncertainty

There were two major sources of quantifiable uncertainty in the above analyses: the interpolated data's agreement with ground-truth data (classification certainty), and the interpolation itself (modelling certainty). Both of these sources were resolved and examined.

Error matrices were used to determine the extent of the classification certainty for each of the maps. The ground-truth data was compared to the interpolated and classified acoustic data via the remaining 80% of the underwater video segments that were not used in the prior analysis. The details of the agreements and disagreements were laid out into an error matrix which tallied the number of occurrences of each of the nine possible outcomes (Table 2.2). These nine outcomes can be consolidated into four types relative to the substrate under consideration, including: true positives (TP), where the ground-truth data and the interpolated datasets agree on the presence of a substrate (ie. V_1A_1 for RK); true negatives (TN), where the datasets agree on the absence of a substrate (ie. when considering RK, V_2A_2, V_3A_2, V_2A_3, and V_3A_3); false positives (FP), or errors of commission, where the interpolation falsely predicts presence (ie. for RK, V_2A_1 and V_3A_1); and false negatives (FN), or errors of omission, where the interpolation falsely predicts absence (ie. for RK, V_1A_2 and V_1A_3). These outcomes were used to calculate five statistics for further investigation (Table 2.3). These statistics are highly correlated due to the fact that they are all derived from the outcomes of the error matrix, but they are useful in determining the relative accuracy of classification for each substrate category.

The modelling certainty was described visually. In the creation of the maps, the substrate category for each cell was selected by choosing the category with the highest probability and this probability represented the certainty in the interpolation. The certainty was displayed as a map layer with the raster cells coloured to represent their percent certainty.

These two analyses of uncertainty were used in tandem to select which of the six substrate maps were the most accurate and therefore which sets to further analyse. As the definition of "most accurate" is exceptionally subjective and entirely reliant on the goals of both the map-maker and the end user, the details of this decision are discussed at length later in this document. The two substrate maps selected were *100-5* (100 m track spacing, 2.5 m spatial resolution (5 pings, 0.5 m apart)) and *50-20* (50 m track spacing, 10 m spatial resolution (20 pings, 0.5 m apart)).

Table 2.2: A) The assignment of each substrate category in the two data sets to a placeholder label. B) An error matrix representing the nine possible outcomes of the comparison between the acoustic interpolation and the video segments.

A.

Video	Acoustic
V_1 = RK	A_1 = RK
V_2 = SA	A_2 = SA
V_3 = GV	A_3 = GV

B.

Acoustic Interpolation (Prediction)	Video (Ground-truth)		
	V_1	V_2	V_3
A_1	V_1A_1	V_2A_1	V_3A_1
A_2	V_1A_2	V_2A_2	V_3A_2
A_3	V_1A_3	V_2A_3	V_3A_3

2.2.4.1 Trap Data

Using a combination of R v3.4.1 and ArcMap v10.5, the 2089 photos that comprised the lobster trap data were reduced in number. This included removing those photos taken over land, at an oblique angle (ie. gimbal pitch <88°), from an altitude of lower than 85 m, or not facing within 10° of north. The photos were then further reduced by eliminating those that were redundant due to complete overlap. This left 168 photos efficiently covering the entire area flown by the drone. These photos were then visually inspected for lobster trap buoys. A corresponding data point was created and georeferenced at each buoy. If the rope connecting the buoy to the trap was visible through the water (Figure 2.3), the point was placed at the trap end to increase accuracy. This created a dataset of lobster presence data that was later used in the presence-only maximum entropy analysis.

Table 2.3: Descriptions and equations of the statistics derived from the error matrices (adapted from *Barrell et al.* (2015))

Statistic	Equation	Description
Overall Accuracy	$\frac{TP+TN}{n}$	Proportion of all predictions that were correct
Sensitivity	$\frac{TP}{TP+FN}$	Proportion of correctly predicted presences
Specificity	$\frac{TN}{TN+FP}$	Proportion of correctly predicted absences
Positive Predictive Value (PPV)	$\frac{TP}{TP+FP}$	Proportion of positive predictions that are TP
Negative Prediction Value (NPV)	$\frac{TN}{TN+FN}$	Proportion of negative predictions that are TN

2.2.5 MaxEnt

The two substrate maps selected after the certainty investigation (*100-5* and *50-20*) were paired with the presence-only lobster trap data and LLWLT tidally-corrected depth rasters and processed through MaxEnt, a species distribution and environmental niche modelling software (*Phillips et al.*, 2006). MaxEnt requires georeferenced presence-only species data (ie. the lobster trap points) and independent environmental variable rasters (ie. substrate maps and depth rasters). MaxEnt was used to develop four SDMs based on the data, displaying each model in two different output formats. MaxEnt also created receiver operating characteristic (ROC) curves, which plot the True Positive Rate (TPR) against

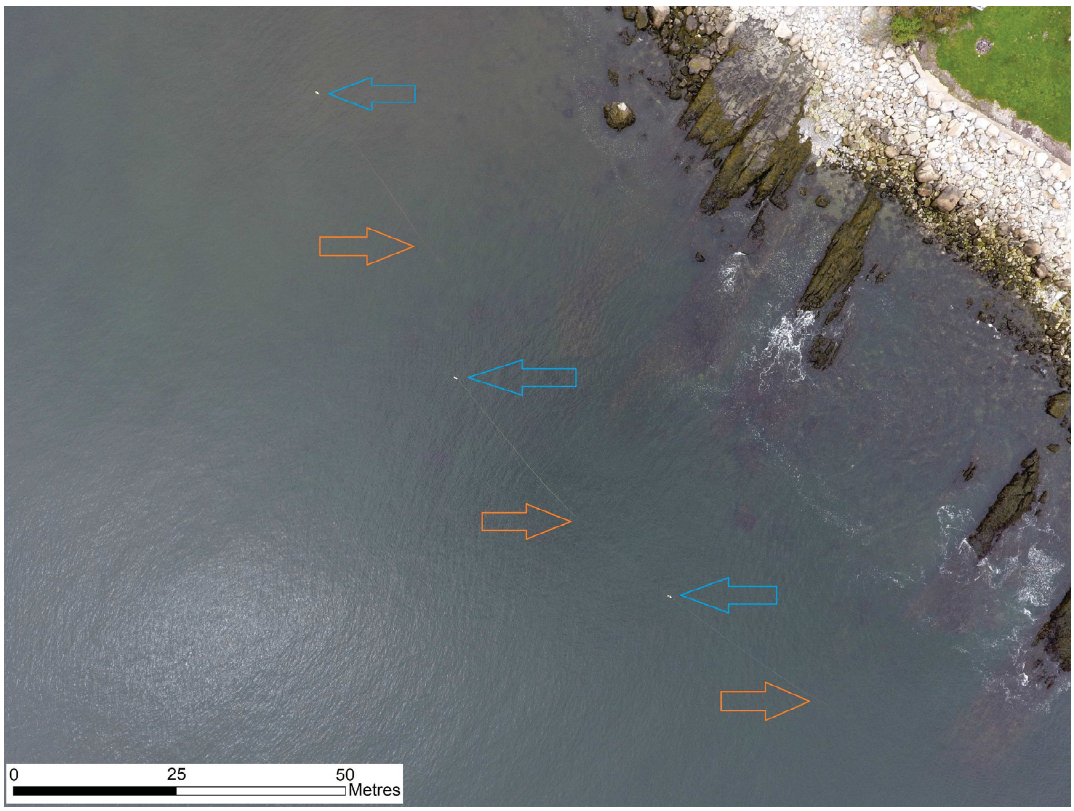

Figure 2.3: An aerial photograph of three lobster trap buoys in Liverpool Bay. The blue arrows indicate the buoys floating on the surface of the water, and the orange arrows indicate approximately where the trap is located on the end of the visible rope.

the False Positive Rate (FPR) to demonstrate the accuracy of the model, and histograms displaying the average species response to the different substrates.

The SDMs produced via MaxEnt come in four possible output formats, two of which were selected for the purposes of this work: "raw" and "cloglog". The raw output unit is relative occurrence rate, or ROR, which is the relative likelihood of species presence in one raster cell as compared to another (*Phillips et al.*, 2006). The unit for the cloglog output is probability of presence, or POP, which is the absolute likelihood of species presence within a raster cell. For more detail on MaxEnt and its outputs, refer to **Appendix A**.

2.3 Results

2.3.1 Substrate Maps

Rather than explain results for all six data sets, this Results section focuses on the two substrate maps selected for further processing in MaxEnt, *100-5* and *50-20*. The rationale behind this selection is explored in Section 2.3.2.

2.3.1.1 100-5

The substrate classification resulted in 9148 MPs classified as rock (RK), 19,620 as sand (SA), and 10,806 as gravel (GV), for a grand total of 39,574 MPs (Figure 2.4). The majority of the SA MPs are located in the central channel portion of the bay, with RK and GV to the north and south. RK MPs are generally found along the coastal edges of the surveyed area, notably near Coffin Island and the northernmost edge of the bay. The bulk of the GV MPs are between the RK areas and the SA, particularly southwest of the island. The northern half of the surveyed area appears to be more heterogeneous, with the RK, GV, and SA MPs intermixed, than the middle SA-heavy section.

The interpolation of the MPs produced a map with a number of small "islands" of substrate surrounded by a different substrate category, such as the two points of RK in the middle of the northeast section of GV and the patches of SA in the northwest swath of RK (Figure 2.5). The central channel SA majority that was shown in Figure 2.4 appears in the interpolation, with most of the southern half of the survey area predicted as SA. The northeastern extreme of the map, which extends to Coffin Island, predicts a strip of RK that then gives way to GV further extending to approximately half of the northern section. Notably, this section of GV prediction is located adjacent to the passage between Coffin

Island and the mainland and is further from land than the neighbouring sections where RK was predicted, which are generally near the shore.

2.3.1.2 50-20

Substrate classification resulted in 2535 MPs classified as RK, 5365 as SA, and 1951 as GV, totalling to 9851 MPs (Figure 2.4). The classifications are less homogeneous in comparison to those of *100-5*, with the northern portion of the map containing more SA MPs and the central channel area of the bay displaying a less clear trend towards SA than is visible in *100-5*. There are four tracks in the central channel section with segments that are heavily classified as RK, in contrast to the tracks surrounding them, and are likely an artifact. The southernmost edge of the surveyed area has a trend towards GV, as does the northeastern edge before it turns into RK MPs near Coffin Island; these two sections are both similar to what is displayed in the *100-5* map.

Despite the diminished trend towards SA in the central channel, the interpolation still predicts SA throughout the entire middle section of the map (Figure 2.5). SA dominates this interpolation, encompassing much of the northern half of the map in addition to the central channel. The section adjacent to the passage between Coffin Island and the mainland is predicted as majority SA, with some patches of both RK and GV within it. The strip of RK and then GV by Coffin Island is predicted in a similar manner to that in *100-5*, but the GV transitions into SA at a much closer point to the island. RK is absent from the southern half of the predictions, occurring only near the shore on the northern edge and near Coffin Island.

The variation visible in the non-interpolated data, particularly in the central channel and the southern shore, was largely made uniform by the interpolation and this was likely due to the smoothing factor. Since kriging tends to create artificially sharp boundaries between neighbourhoods, I applied a smoothing factor to reduce the effect of the mathematical artifacts. However, the smoothing also removed some of the variation in the data, creating unexpectedly homogeneous predictions. While the smoothing was done in both *100-5* and *50-20*, it is more evident in *50-20* due to the higher variance in the non-interpolated data.

2.3.1.3 Depth

The 1 m depth contours show a shallow gradient of increasing depth in the west to east direction through the bulk of the survey area, ranging from 4.3 m to 29.2 m deep

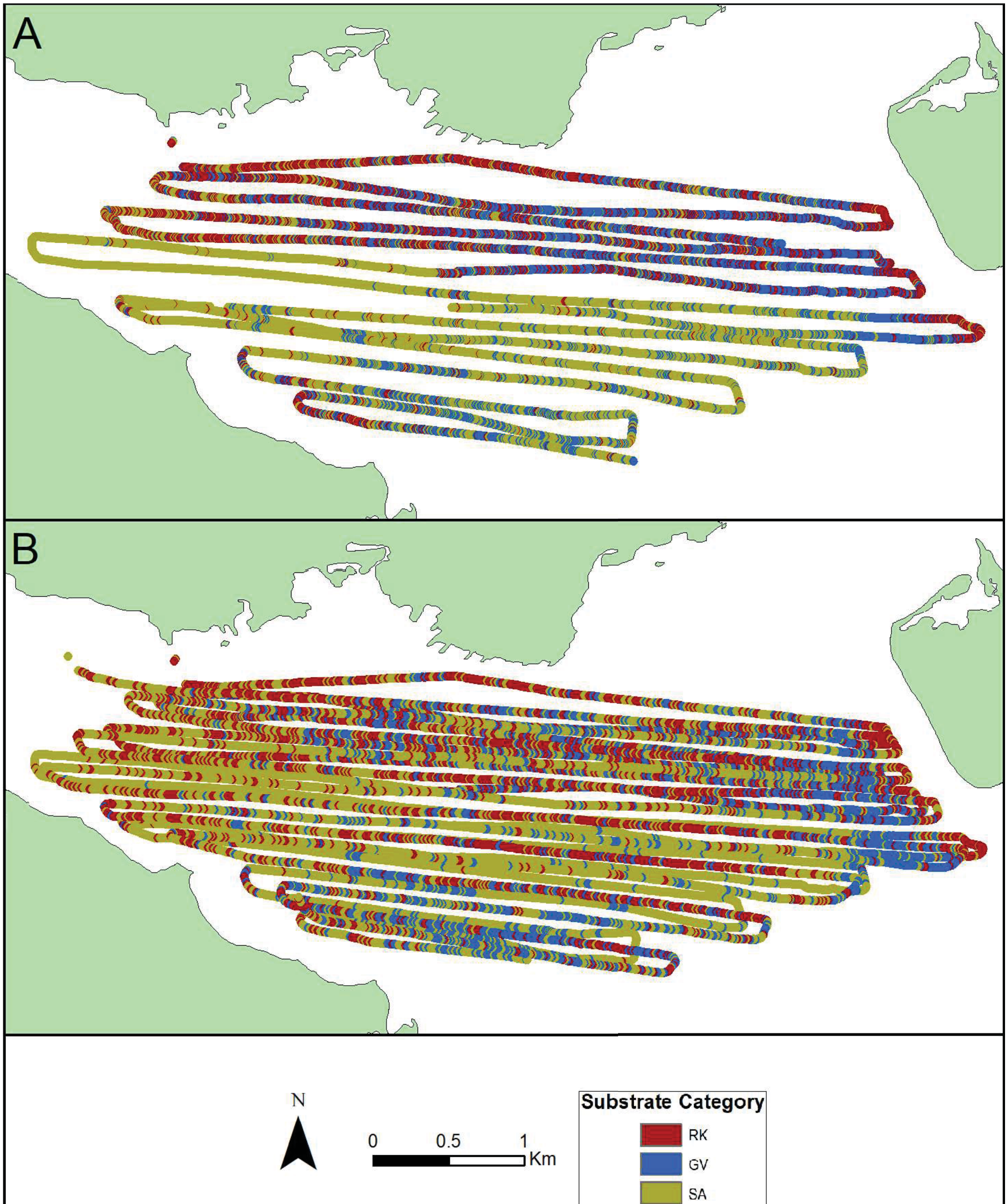

Figure 2.4: Acoustic data tracks with MPs classified into substrate categories; A: *100-5*, B: *50-20*.

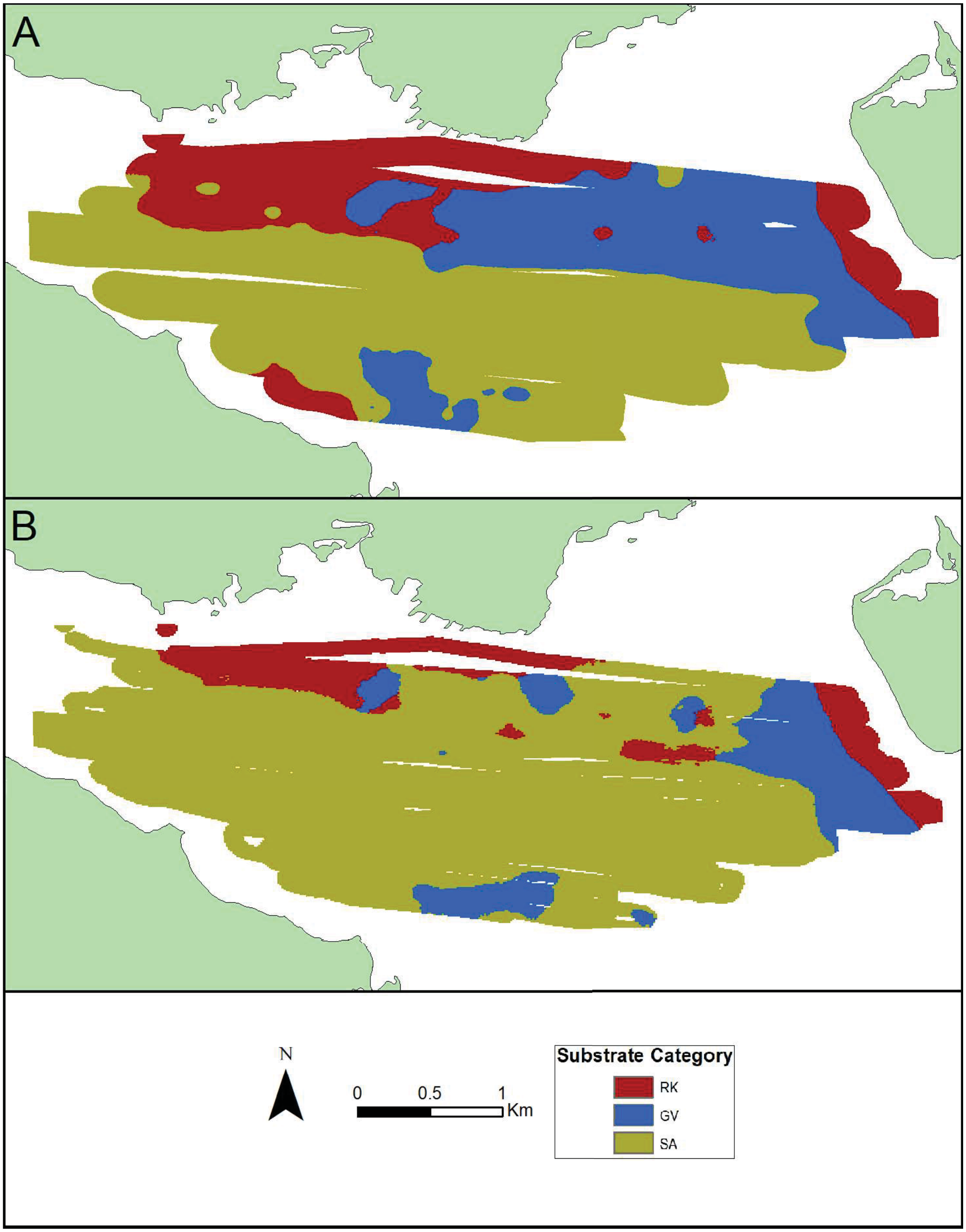

Figure 2.5: Interpolated data, completed via indicator kriging, showing the distribution of substrate category predictions; A: *100-5*, B: *50-20*.

(Figure 2.6). Near Coffin Island, the gradient is significantly steeper than the central portion and it is relatively uniform in comparison to the gradient on the northern shore. These steep gradient areas to the north and near the island are also where both *100-5* and *50-20* show distribution of RK and GV. The southern shore displays a slightly steeper gradient than the central channel, which corresponds to RK and GV in the *100-5* substrate map.

2.3.2 Validation and Uncertainty

Selection of *100-5* and *50-20* data sets for further analysis was based on several factors. *50-20* had the highest RK accuracy, one of the highest overall classification accuracies, and was one of the best at correctly predicting where RK would be (high PPV value). Considering lobsters' distinct preference for complex rocky substrates, this ability to distinguish between soft sediments and rock is important. *100-5* was similar to *50-20* in the highest RK PPV value, despite the low overall classification accuracy of the 2.5 m resolution maps.

2.3.2.1 100-5

The classification error matrix and statistics for *100-5* (Tables 2.4, 2.5) showed that the RK and GV categories were frequently confounded, with RK predicted as GV for 43.59% of the predictions. This model had a tendency to err on the side of RK or GV, meaning that it was ~3x more likely that SA points were incorrectly predicted as either RK or GV compared to RK or GV points incorrectly predicted as SA. For example, the sensitivity value for SA, which is the number of actual SA points that were correctly predicted as SA, is 8.57%, a much lower value than the sensitivity of RK (38.46%) or GV (35.29%), indicating that many actual SA points were predicted to be another category. Most of the classification errors (FN and FP) are due to these two trends.

The PPV for RK, which is the percentage of positive RK predictions that were TP, is 42.85% (SA 16.67%, GV 15.79%). This suggests that this model is best at finding RK in comparison to the other two categories. The overall accuracy for each of the substrate categories in this map is approximately 50%. For all three categories, the overall accuracy was more strongly affected by TNs than by TPs.

The modelling certainty map showed high certainty in the centre of the bay and along most of the extreme edges of the survey area near the shore (Figure 2.7), indicating that

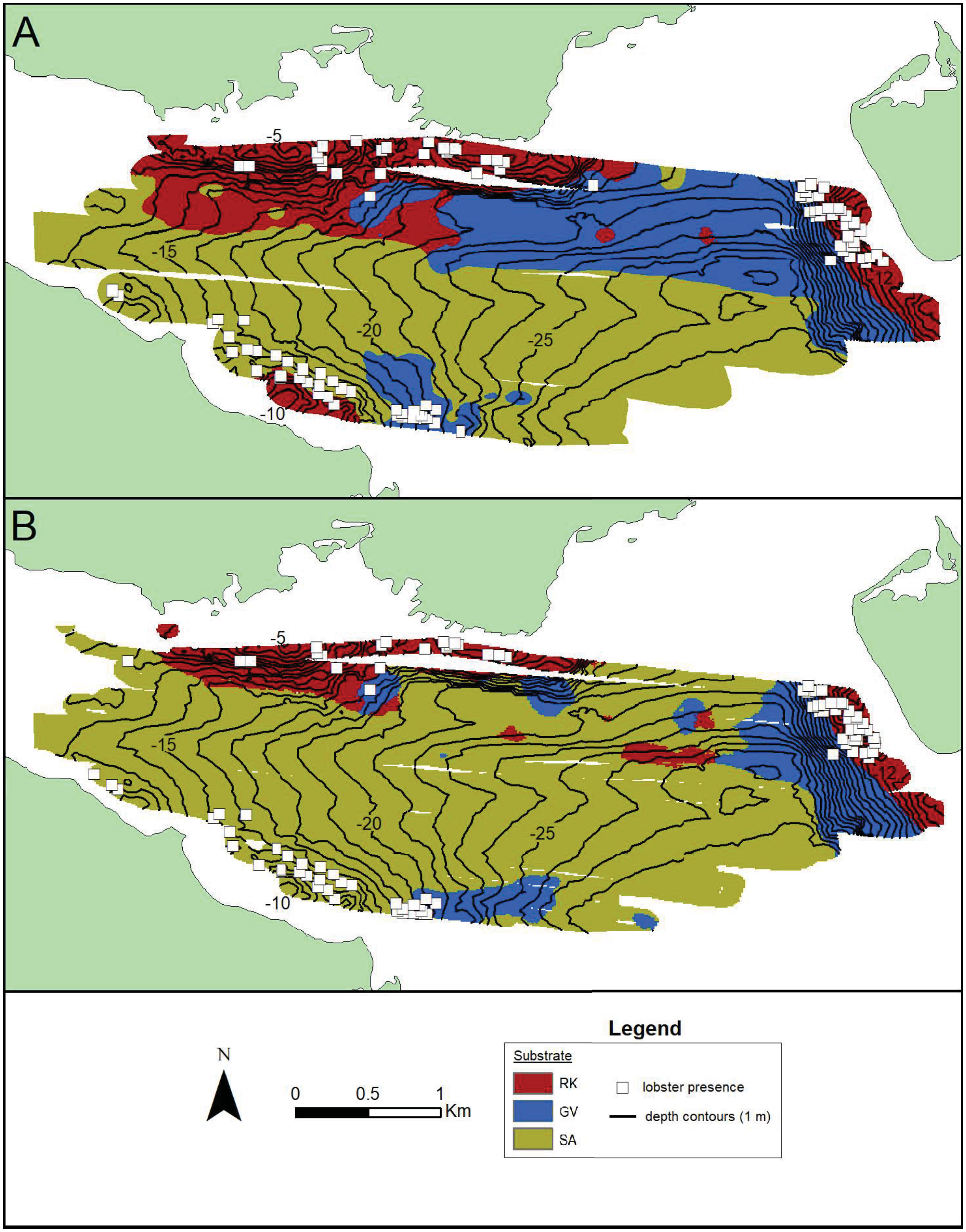

Figure 2.6: Substrate distribution maps for both A: *100-5* and B: *50-20* with 1 m depth contours and lobster presence points.

the substrate is more homogeneous in these locations. The northern third of the map had a lower certainty, as did a small section to the south. Lower certainty coincided with higher heterogeneity (Figure 2.4). The vast majority (34 of 44) of the RK classification errors uncovered in the matrix are located within the southern patch of low certainty.

2.3.2.2 50-20

The error matrix and statistics for *50-20* (Tables 2.6, 2.7) showed an opposite trend compared to *100-5*, where instead of erring on the side of the more complex substrate, it erred on the side of SA. Both RK and GV were $\sim$ 3.3x more likely to be incorrectly predicted as SA than the other way around. The sensitivity of SA is 68.57%, compared to RK at 16.21% and GV at 28.57%, meaning that there were a much higher number of SA points correctly predicted as SA than RK as RK or GV as GV. Similarly, the specificity (or proportion of correctly predicted absences) is relatively high for RK (83.67%) and GV(90.27%) in comparison to SA (27.45%); SA is over-predicted for presences and therefore results in fewer correctly predicted absences. This over-prediction is visually depicted in the difference in variation between the non-interpolated data, which shows heterogeneity across much of the map, and the interpolated data, which displays large sections of homogeneous SA.

Table 2.4: Error matrix for *100-5* representing the comparison between the acoustic interpolation predictions and the underwater video.

Acoustic Interpolation (Prediction)	Video (Ground-truth)		
	RK	SA	GV
RK	15	17	3
SA	7	3	8
GV	17	15	6

Table 2.5: The error statistics for the *100-5* map as a function of substrate type.

	Overall Accuracy	Sensitivity	Specificity	PPV	NPV
RK	0.5165	0.3846	0.6154	0.4286	0.5714
SA	0.4835	0.0857	0.7321	0.1667	0.5616
GV	0.5275	0.3529	0.5676	0.1579	0.7925

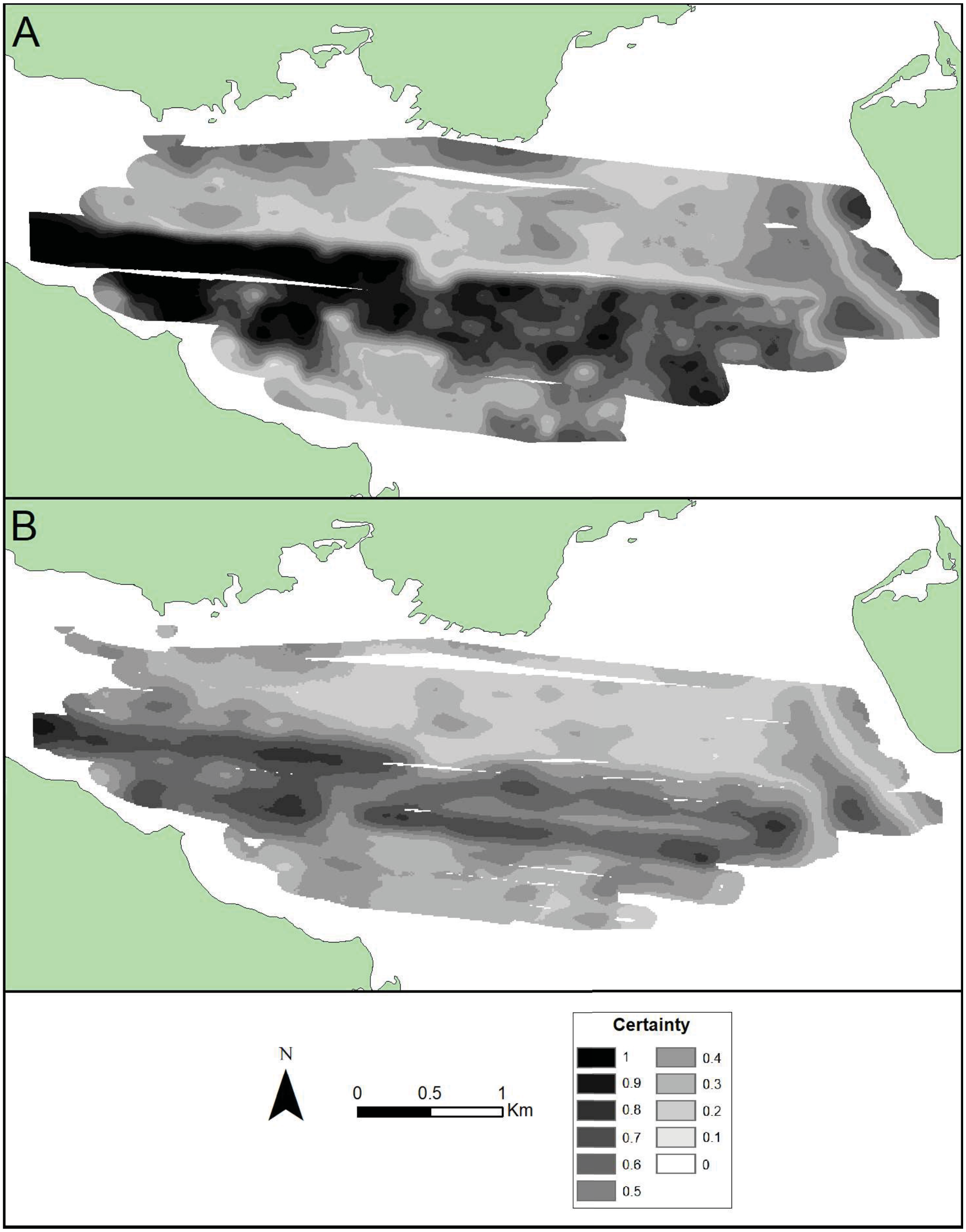

Figure 2.7: Modelling certainty maps derived from the probability remainder of the interpolated data; A: *100-5*, B: *50-20*.

Table 2.6: Error matrix for *50-20* representing the comparison between the acoustic interpolation predictions and the underwater video.

Acoustic Interpolation (Prediction)	Video (Ground-truth)		
	RK	SA	GV
RK	6	8	0
SA	27	24	10
GV	4	3	4

Table 2.7: The error statistics for the *50-20* as a function of substrate type.

	Overall Accuracy	Sensitivity	Specificity	PPV	NPV
RK	0.5465	0.1622	0.8367	0.4286	0.5694
SA	0.4419	0.6857	0.2745	0.3934	0.5600
GV	0.8023	0.2857	0.9028	0.3636	0.8667

The PPV for RK is the same as for *100-5* at 42.85%. While it is higher than the SA and GV categories (39.34% and 36.36%, respectively), the difference is not substantial. This indicates that the ability of this model to distinguish one category over another is approximately equal among the categories. The overall accuracies for the categories were 54.65% for RK, 44.18% for SA, and 80.23% for GV, all of which were strongly influenced by TNs.

The modelling certainty map is similar in structure to that of *100-5* but with a trend of lower certainty throughout the entire bay (Figure 2.7). The northern portion of the bay and a section on the southern shore show low certainty, which coincided with higher heterogeneity (Figure 2.4).

2.3.3 MaxEnt SDMs

2.3.3.1 100-5

The areas under the curve (AUC) of the receiver operating characteristics (ROC) curves are identical for both *100-5* model output formats (Figure 2.8) at 0.846. This is a measure of how well the model performs and is compared to the results of a theoretical model based on random distribution predictions (AUC = 0.5).

The species response graphs, which plot ROR or POP against the variable in question, illustrate a correlation between depth and substrate category (Figure 2.9). The substantial change in response when depth was excluded as a secondary variable indicates that depth and substrate do not occur independently of one another. Depth has a more significant effect on the overall species response than does substrate category. This is confirmed in the MaxEnt environmental variable contribution statistics, which demonstrate that depth contributed 85.7% to the model and substrate category only 14.3%. For both the raw and the cloglog output models, the response to substrate category (with depth excluded) is strongest in RK (respectively: 2.42x10^{-4}, 0.847), followed by GV (9.40x10^{-5}, 0.518), and SA (4.38x10^{-5}, 0.288).

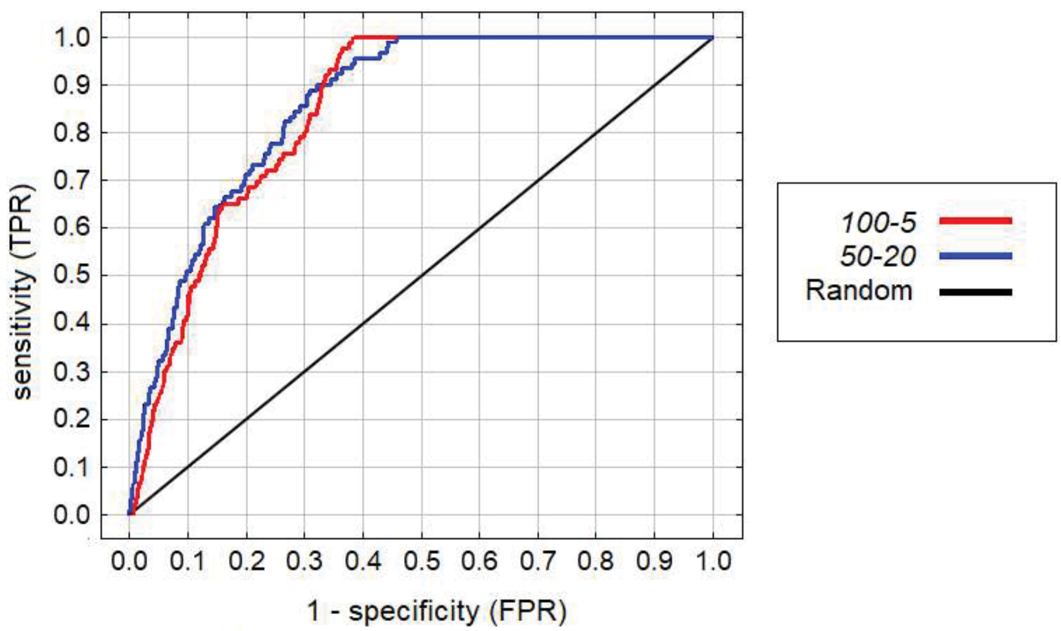

Figure 2.8: ROC curves for both *100-5* and *50-20*. Curves that exist above the theoretical model at 1:1 (AUC = 0.5) are considered better than random. *100-5* has an AUC of 0.846 and *50-20* has an AUC of 0.860.

The raw output shows high RORs along the shorelines, particularly against Coffin Island (Figure 2.10). The bulk of the centre and the southeast corner contain substantially lower RORs, particularly as the bay opens to the ocean. The RORs along the southern edge, near the shore, are generally lower than those at the other two clusters of lobster traps (ie. the northern and Coffin Island shores). The western section of the central channel has relatively high RORs given the absence of lobster traps.

The cloglog model output predictions are similar to those of the raw model, with high POP values in the same areas (ie. coastal, particularly along Coffin Island) and low POP

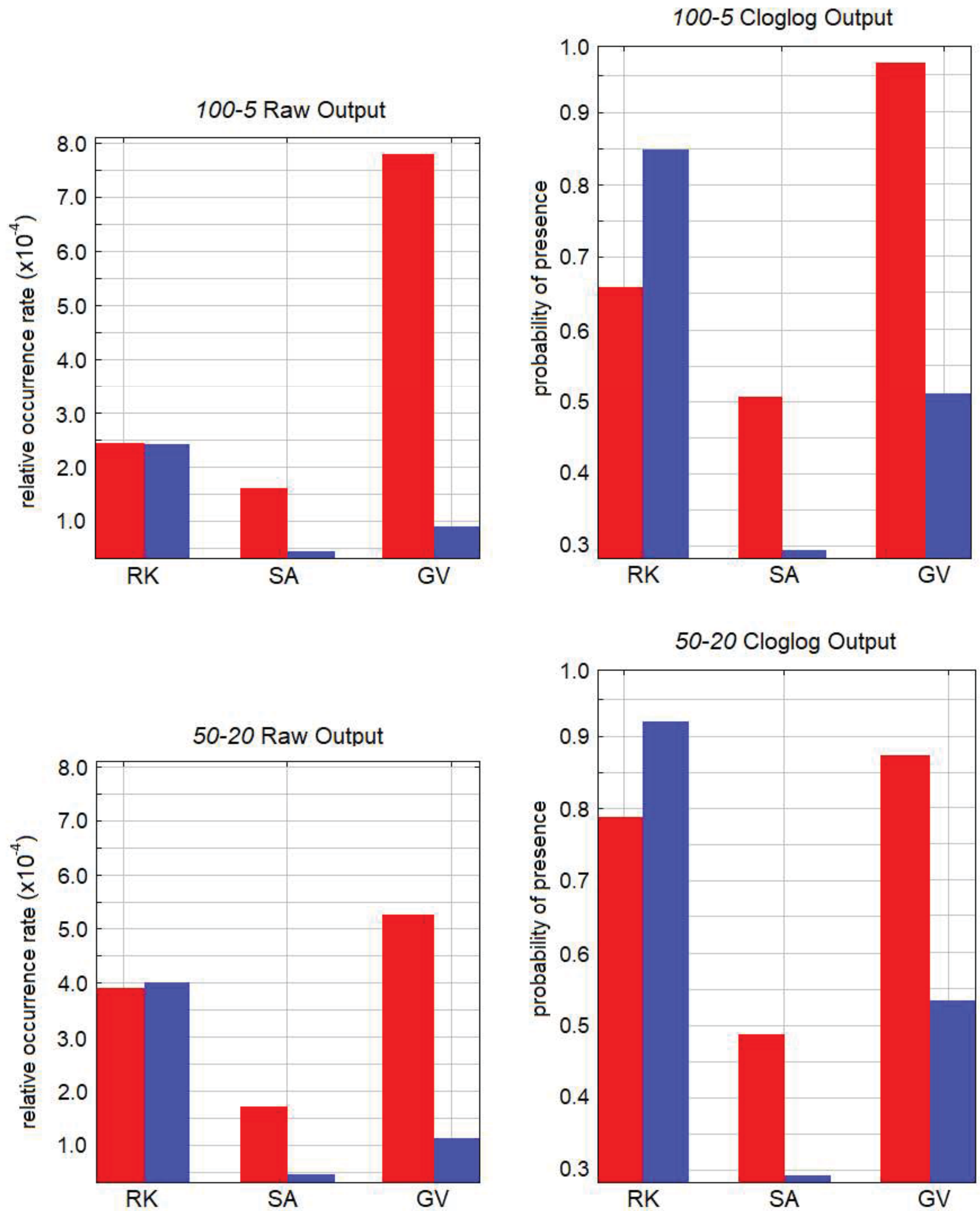

Figure 2.9: Species response to substrate categories. The red is the marginal response to the substrate variable with the depth variable included; blue is the response to the substrate variable without the depth variable.

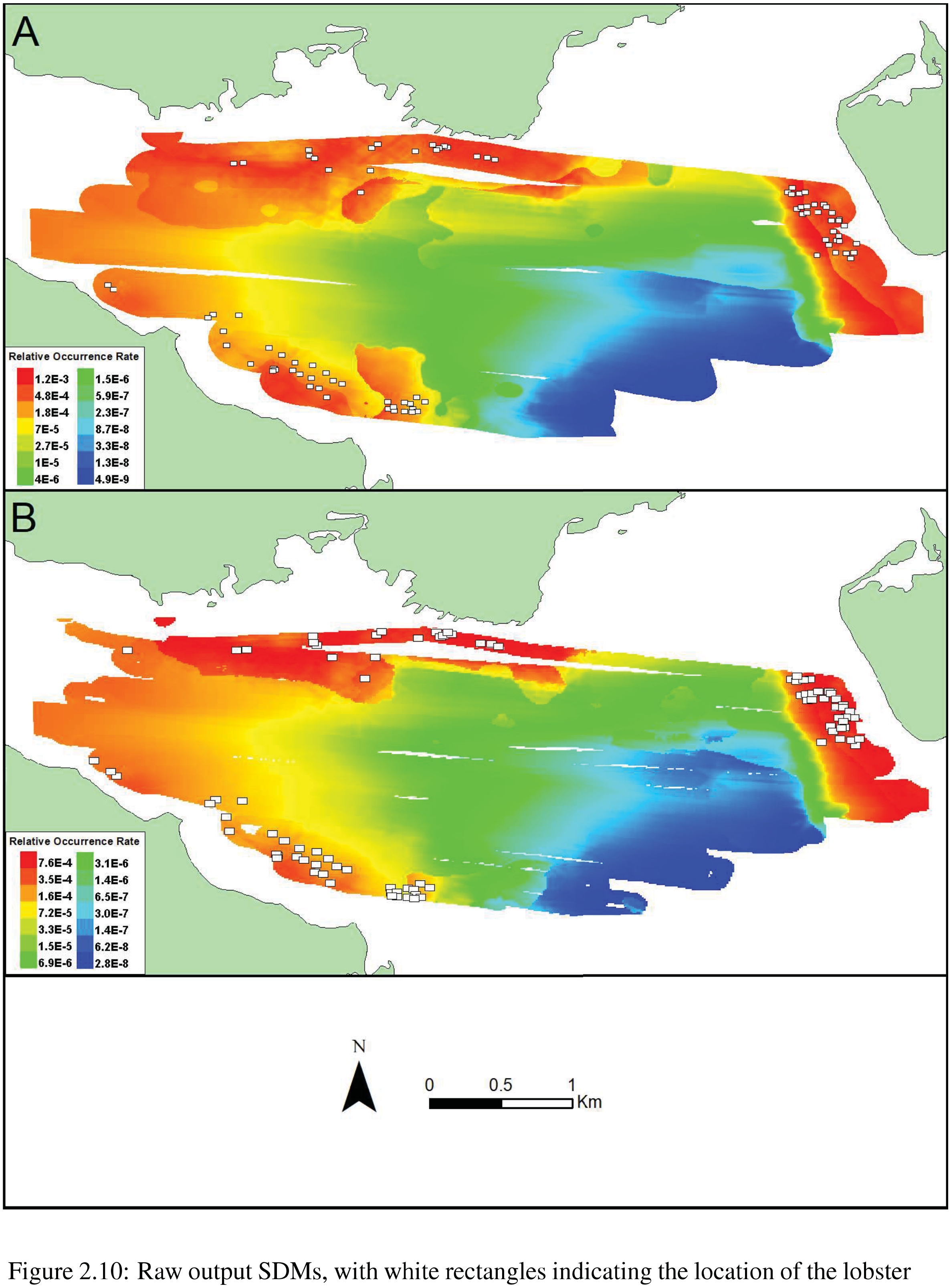

Figure 2.10: Raw output SDMs, with white rectangles indicating the location of the lobster traps; A: *100-5*, B: *50-20*. Note the difference in the colour scales.

values in the central and southeastern sections (Figure 2.11). While not immediately comparable, the cloglog output appears to label more area as less suitable than does the raw model.

2.3.3.2 50-20

Again, the AUC for the ROC curves are identical for both *50-20* model output formats (Figure 2.8) at 0.860. This AUC is not substantially different from the *100-5* AUC value of 0.846.

The species response graphs reveal the same confounding problem as in *100-5*, where the substrate category variable is not independent from the depth variable (Figure 2.9). The dependence is skewed in a similar manner to that of *100-5*, such that the response change is much larger in the substrate category variable when the depth is removed versus the change in the depth variable when the substrate category is removed. This implies that the depth variable has more weight in the model, an idea that is supported by the environmental variable contribution statistics calculating that depth contributed 71.4% to the model and substrate category contributed 28.6%. Again, both the raw and the cloglog outputs responded strongest to substrate (depth excluded) in RK (respectively: $4.00\text{x}10^{-4}$, 0.929), then GV ($1.15\text{x}10^{-4}$, 0.533), and finally SA ($4.97\text{x}10^{-5}$, 0.280).

Much like *100-5*, the raw output for *50-20* displays high ROR values near the three coastlines. Near Coffin Island, the RORs are almost uniformly extremely high. The southern coastline has lower RORs than the equivalent area in the *100-5* raw model. Again, the western section has high RORs considering the absence of traps.

The cloglog output for *50-20* is visibly less patchy in comparison to the cloglog output of *100-5*, with smoother transitions in notable areas (eg. Coffin Island) (Figure 2.11). The northern shore cluster of high POPs is higher in this model than in *100-5*. There are some barely visible patches of higher POP values in the northeastern quadrant of the bay, contrasted against the zero value POPs that encompass most of that area; these patches correspond to GV and RK patches on the substrate map.

2.4 Discussion

While the overall results of the maps created from both *100-5* and *50-20* are similar, both demonstrating higher species responses along Coffin Island, the northern shore, and to

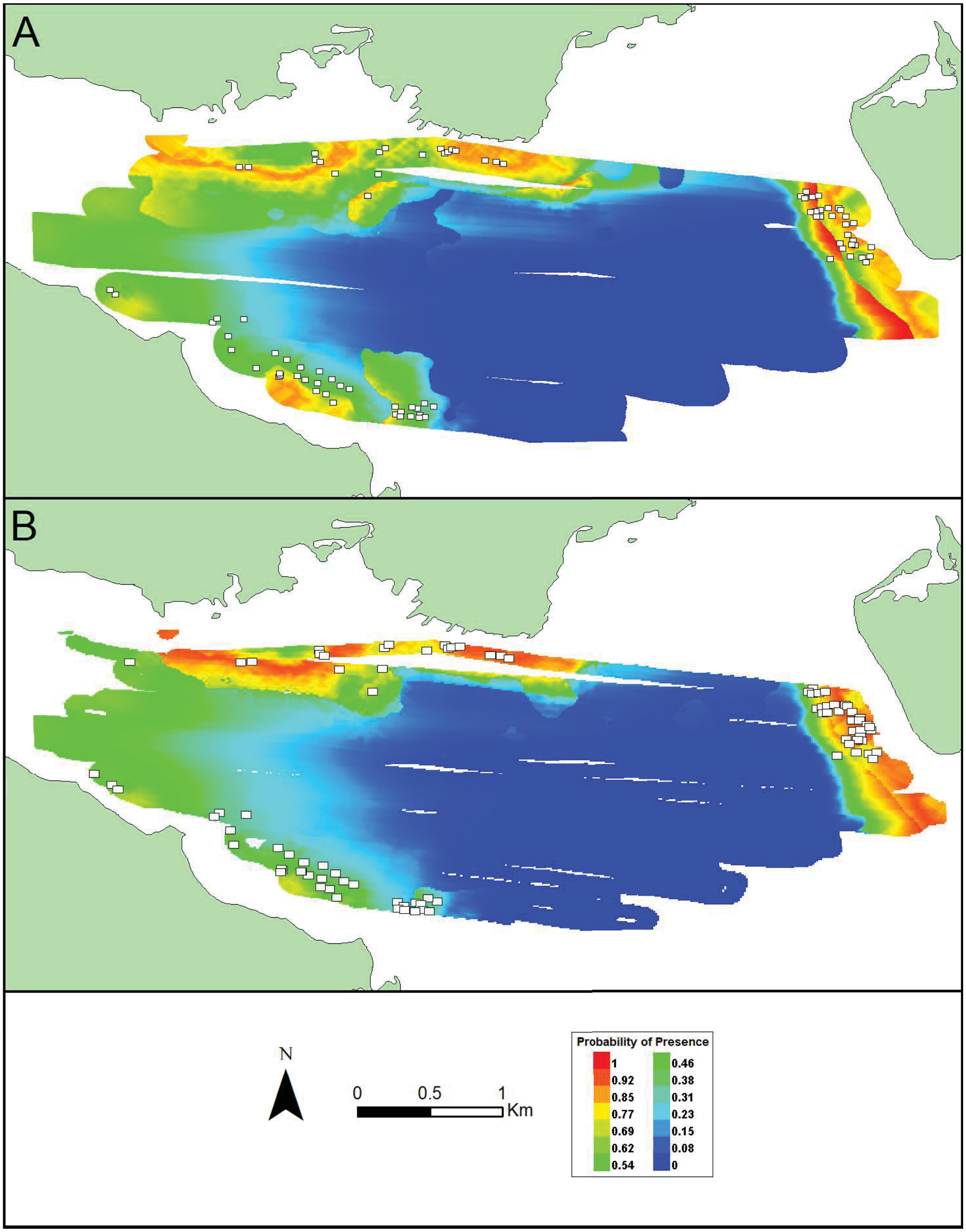

Figure 2.11: Cloglog output SDMs, with white rectangles indicating the location of the lobster traps; A: *100-5*, B: *50-20*.

a lesser extent the southern shore, there are distinct differences between them requiring examination to develop an appropriate interpretation.

2.4.1 Resolution and Track Separation

The creation of six sets of data, an intersection of two track separation distances and three spatial resolutions, came from unknowns in the development stage of this work: what is the optimum combination of point density and point binning? Despite the fact that spatial scale can have strong implications on the outcome of the map layers (*Brown et al.*, 2011), there is no standard for choice of values.

There is a large variety of track spacing found in the literature for surveys of similar scales, depths, and near-shore locations as this one (*Freitas et al.*, 2011; *Anderson et al.*, 2002; *Brown et al.*, 2005; *Foster-Smith et al.*, 2004). Given that the footprints of the acoustic pings collected in Liverpool are very small (0.16 – 23.00 m^2) in comparison to most of the literature examples of track spacing distances (2 km – 70 m), the interpolation between the pings would be sparsely informed if those values were adopted. The relatively small size of Liverpool Bay encouraged the choice of tighter track spacing (*Brown et al.*, 2005). However, effort, which is inversely proportionate to the spacing, also needed to be considered. The track spacings were chosen to balance these two conflicting concepts.

The literature provides discussion of the spatial resolution limitations of optical data (*Schweizer et al.*, 2005), but SB-AGDS studies tend to avoid the matter (*Lecours et al.*, 2015). The selection of too coarse a spatial resolution can lead to what is commonly known as the "mixed pixel problem", where the value of a pixel is the result of a combination of signals from different reflective groups (*Schweizer et al.*, 2005; *Jones and Sirault*, 2014). However, too small a spatial resolution allows for spurious detail, or "noise", to be collected and recorded as valid variation (*White et al.*, 2003).

The effects of the different spatial scales caused the differences in substrate distribution between *100-5* and *50-20* in the north-eastern quadrant. The low modelling accuracy in the area indicates that interpolation was not effective, suggesting high heterogeneity in the substrate. This heterogeneity can lead to differences in binned averages if the bin sizes vary, which is how the pixels of the maps were constructed.

2.4.2 Set Decision

While both classification accuracy and modelling accuracy are crucial in creating these species distribution maps, classification accuracy was favoured compared to the modelling accuracy. Matching the substrate categories to reality is more critical to the model than the interpolation. *Hutin et al.* (2005) and *Freitas et al.* (2006) both focused on the classification and avoided interpolation entirely.

The two maps ended up portraying two extremes of a risk spectrum. As reported in Section 2.3.2, *100-5* errs on the side of RK and GV, which has the risk of overestimating lobster habitat. However, the *50-20* results err on the side of SA, and could be considered to underestimate lobster habitat. These trends are most easily observed in the interpolated substrate maps (Figure 2.5).

2.4.3 MaxEnt and ROCs

The maximum entropy on which MaxEnt is based is a statistical prediction method which assumes the flattest, most uniform probability distribution possible while allowing for constraints, which are usually represented by the measured variable data (*Jaynes*, 1957). In a uniform probability distribution, any single outcome is equally as likely to occur as any other, which leads to a high uncertainty in outcomes, also known as a high entropy (*Harte and Newman*, 2014). The maximum entropy method acknowledges the environmental constraints present while simultaneously maximizing the entropy (ie. smoothing the probability distribution into near uniformity) around them. Using anything other than the maximum entropy in the distribution outside the constraints would require the assumption of distribution information that is unknown.

ROCs which use true absences have a maximum AUC of 1, which indicates that there were no false positive predictions in the model. However, ROCs made with pseudo-absences, such as those in this study, have a maximum AUC that is both less than 1 and unknown. Therefore it is difficult to predict how well a model has done compared to its unknown optimum. Instead, the ROCs presented in this study should be compared to the random model (AUC = 0.5) and to each other.

2.4.4 Depth

Depth occurred as the main contributing variable in both *100-5* (85.7%) and *50-20* (71.4%), meaning that depth had a stronger influence on both the responses of the models than the substrate categories. Depth was included in the modelling process because it is a standard oceanographic variable and is temporally stable at similar scales to substrate distribution. However, the entire depth range of the surveyed area is well within the habitable range for the American lobster (*Holthuis*, 1991), which, in addition to the evidence of correlation revealed in the response histograms (Figure 2.9), suggests that the depth effect in the models is misleading. If depth is a major causative variable, the western section of the surveyed area which is both shallow and sandy (Figure 2.6) would likely have more lobster traps. This is an indication that the depth variable is not the main contributing factor to species response. Instead, the pattern that is detected by the model through the depth variable appears to be more closely related to distance from shore (Figure 2.10, 2.11; the southern shore has high probability despite relatively deep water). Distance from shore is highly correlated with substrate type, as substrates are generally large near shore and smaller with distance, which indicates that the high suitability on the shorelines is likely to be the result of substrate and not depth.

The effect of depth changed the magnitude and ranking of the substrate categories (Figure 2.9). In all four model outputs (raw and cloglog for *100-5* and *50-20*), GV created the most response and RK the second most when depth was included, but those positions reversed when depth was removed from the analysis; SA remained as the least responsive category, though with a change in magnitude following the removal of depth. Areas categorised as GV span a wide range of depths, from approximately -8 metres to -27 metres, and most of the lobster presence points found in GV areas are located in the shallower sections, particularly along the shore of Coffin Island. Therefore, the relationship between the species response and GV when depth is included is expected: shallow GV has a high species presence. When depth is removed as a variable, all of the GV areas are considered equally and since there are very few presence points on the large patches of deeper GV, the species response averages out to values much lower than they were when depth was included. The lack of lobster presence points in deep gravel areas is likely due to the increased distance from RK areas, which is the more preferred substrate. Because there is less variance in depth in the areas categorised as RK (roughly -4 metres to -14 metres), the removal of depth has a lesser but opposite effect. Since shallow depths were associated with high presence (due to both the GV-depth relationship described above and the shallow, high presence nature of RK), the removal of depth meant that the deeper RK areas no longer had a negative effect on the overall response, leading to an increase. Despite this explanation of the interdependency between depth and substrate, removing depth from consideration entirely and using only substrate to predict species occurrence would be inappropriate. Even when confounded with other variables, depth has a substantial effect on habitat suitability because the best lobster habitats tend to be found in shallower water (*Lawton and Lavalli*, 1995).

2.4.5 Substrate Categories

Despite the confounding nature of depth on substrate category, the switch in response rankings between RK and GV in reaction to the exclusion of the depth variable does not substantially alter the final map interpretation. Descriptive resolution is a term first coined by *Green et al.* (1996) to encompass the varying breadth and detail of substrate category descriptions revealed by remote sensors. The descriptive resolution of the SB-AGDS in this study is quite coarse, revealing the substrate categories rock/RK, gravel/GV, and sand/SA. This coarseness has created broad substrate categories and given that RK and GV are relatively similar in physical structure, especially when contrasted to SA, there is a high probability that a) they share some characteristic features, and b) they were often mistaken as each other during the cluster analysis, before the interpolation stage. This is, in essence, a more abstract version of the mixed pixel problem. If the SB-AGDS were able to define narrower substrate categories, the assignment of substrates into those categories would likely be more precise, have less potential overlap, and display a clearer relationship between substrate and lobster response.

Differences in descriptive resolution were the rationale behind altering the substrate categories from what *Tremblay et al.* (2009) used (Table 2.1) – the video system in their study was able to detect the difference between boulders and cobbles, but my acoustic system was not. *Tremblay et al.* (2009) found fewer lobsters on cobble substrate than among boulders, meaning that the combination of the two categories in this study could have skewed predictions depending on the actual ratio of cobble to boulder in the RK areas. Similarly, *Tremblay et al.* (2009) found very few lobsters on gravel substrate, and in our study the relatively high response may be a result of the proximity of GV to RK (Figures 2.10, 2.11; particularly, see near Coffin Island).

2.4.6 Lobster Traps

Collecting precisely geolocated lobster presence data in Nova Scotia was challenging. A scientific trapping license is difficult to obtain, as are GPS records for traps deployed by fishing boats. No records of geolocated lobsters were found for Liverpool Bay at a suitable scale; government records of landings are binned in grid cells of approximately 18.5 km x 18.5 km (*Coffen-Smout et al.*, 2013). Accompanying a lobster fisher on their boat as they retrieved traps would have required significant financial compensation for their time, which was not available. Additionally, lobster fishers can be extremely protective of their trap placements, as the knowledge of prime trapping spots is valuable (*Acheson*, 2003; a Maine-based reference, but consistent with anecdotal knowledge of local fisheries). However, given the level of detailed knowledge on lobster behaviour and preferred habitat that most lobster fishers have (*Acheson*, 2003), it is reasonable to assume that the presence of a trap buoy corresponds to lobster presence. Therefore, the strength of the local fishers' knowledge of lobster movement was considered reliable enough that the traps they deployed could be used as presence indicators in the model.

One of the major assumptions in using MaxEnt is that the modelled area has been thoroughly sampled (*Elith et al.*, 2010). This assumption was met in that the fishers have sampled the modelled area over an extended temporal scale, using decades of experimentation, trial and error, and shared information. Some change in the location and borders of the known lobster habitat is to be expected over so long a period, but Maxent's boundary modelling is already limited in its precision. However, alternate rationales for trap placement need to be considered. Dangerous or limited access, vessel thoroughfare pathways, debris or wrecks, or local rules and expectations of trapping could all skew the number and placement of traps in a way that is not directly a result of substrate (*Acheson*, 2003).

2.4.7 Application

To avoid user-user conflict between the lobster fishery and the salmon aquaculture industry in Liverpool, new salmon pens could be placed near the centre of the surveyed area, particularly in the green sections of the SDMs. This area is not particularly suitable for lobsters and would satisfy the basic requirements of the net pens, which include shelter from the open ocean and at least 15 m depth. However, other aspects of coastal activity would need to be considered as it is possible that that area has high traffic or experiences other environmental concerns that would negatively affect its suitability to aquaculture.

2.5 Conclusions

This study has shown that the use of SB-AGDS is effective in creating fine scale lobster habitat suitability maps, despite a number of caveats for their use and interpretation. The vague boundaries between substrate patches and the relative nature of the raw suitability distribution values suggest that the maps would be best used as qualitative models, with scales of high and low suitability. The difficulties in obtaining lobster presence data remain, but could be mitigated with the formation of strong relationships between local fishers and researchers. Because of their fine scale, local relevance, and species specific design, the maps created using the above described method are suitable for use as the data layers required in an MSP designed to manage the user-user conflict between the local lobster fishery and the net-pen aquaculture industry. I therefore propose that this method could be used in similar future bay-scale marine spatial planning ventures.

This study also supports the argument that scale is a critical aspect of benthic habitat mapping studies and that the effects of scale and resolution should be explicitly discussed and analysed. The dramatic differences in the two final maps is a direct result of experimenting with spatial scale, and serves as an important reminder to interpret spatial models and map layers with a discerning eye.

CHAPTER 3

HABITAT MAPPING OF JUVENILE AMERICAN LOBSTERS IN LOW COBBLE AREAS USING ACOUSTIC METHODS

3.1 Introduction

Since coastal areas tend to experience heavy anthropogenic use through activities such as fishing, aquaculture, transportation, resource extraction, shipping, and recreation, the spatial organisation of these activities is important to allow compatible use. Marine spatial planning (MSP) is a planning framework that can be defined as the "rational organization of the use of marine space and the interactions between its uses, to balance demands for development with the need to protect the environment, and to achieve social and economic objectives in an open and planned way" (*Douvere*, 2008). As MSP increases in popularity as a management tool in both international and local governments (*Shucksmith and Kelly*, 2014), a deep and specific understanding of the intricacies and characteristics of the spatial data used to construct these individual plans becomes more critical. Spatial habitat data in particular can be challenging to accurately define for a given area as the habitat of a species can vary widely depending on a number of factors, including the interpretation and definition of the term "habitat" (*Howard and Larson*, 1985; *Sly and Busch*, 2018; *Davies et al.*, 2004; *Brown et al.*, 2011; *Laurel et al.*, 2004). Healthy coastal fisheries require sufficient habitat for all life history stages, and for benthic species, initial settlement habitat may be different than adult habitat.

The American lobster (*Homarus americanus*) fishery is the most valuable catch in Atlantic Canada (*DFO*, 2016). Lobsters demonstrate differing preferences in substrate depending on life stage (*Botero and Atema*, 1982; *Wahle*, 1992; *Spanier*, 1993; *Karnofsky and Price*, 1989) and substrate availability (*Tang et al.*, 2015; *Steneck*, 2006), and as such their habitat can fluctuate. Therefore any MSP programs that involve American lobsters will need to address the details of life stage and substrate availability in the data collected in order to be appropriately effective.

Spatial overlap of the lobster fishery with activities such as the net-pen Atlantic salmon (*Salmo salar*) aquaculture industry can cause conflict, particularly concerning lobster population health and trapping access to lobster habitat (*Wiber et al.*, 2012). The delineation of lobster habitat may help avoid these user-user problems and, as demonstrated in **Chapter 2**, lobster habitat can be defined for MSP when a broad range of substrate grain sizes is considered. However, studies have suggested that juvenile lobsters, once thought to settle almost exclusively into cobble (*Palma et al.*, 1998; *Wahle and Steneck*, 1991; *Pottle and Elner*, 1982; *Cobb et al.*, 1983), will settle in large numbers onto less preferred, smaller grain-size substrates if the competition for cobble is high (*Tang et al.*, 2015; *Steneck*, 2006). This introduces a need to examine our ability to produce habitat suitability maps for juveniles in areas of low cobble in order to accurately adopt MSP.

In areas with plentiful cobble substrate, young lobsters settle in their planktonic larval stage into complex, shelter-forming cobble to moult into benthic life stages (*Botero and Atema*, 1982; *Barshaw et al.*, 1994). If larval lobsters encounter less preferred substrate such a mud or sand, they will return to the water column, delaying metamorphosis until finding more suitable substrate (*Botero and Atema*, 1982; *Palma et al.*, 1998; *Cobb et al.*, 1983). Juvenile lobsters seldom live on featureless sediment provided there is substantial cobble nearby (*Wahle and Steneck*, 1991; *Cobb et al.*, 1983). However, this preference for cobble can lead to crowding and competition if cobble is rare. Lobsters are aggressive, solitary animals and high density populations encourage growth inhibition (*Nelson et al.*, 1980; *Cobb and Tamm*, 1974), injury (*Aiken and Waddy*, 1986; *Paille et al.*, 2002), and death (*VanOlst et al.*, 1975; *Paille et al.*, 2002). Small lobsters are more likely to be evicted by large lobsters from cobble substrate (*Steneck*, 2006), resulting in juveniles inhabiting poor quality substrates.

In laboratory conditions, larval lobsters delay their moulting to the first benthic stage when presented with mud substrate, but that delay was significantly longer when only sand was available (*Botero and Atema*, 1982). This indicates that mud is preferred over sand as a substrate on which to settle, yet few studies have reported this difference in preference in the field as sand and mud are frequently grouped together as a single non-cobble substrate (*Wahle and Steneck*, 1991; *Hudon*, 1987). However, it has been shown that juvenile lobsters burrow upon settlement. Juveniles in naturalistic laboratory conditions dug straight burrows beneath stones in cobble substrates and U-shaped burrows in mud (*Botero and Atema*, 1982); if the mud had stones scattered on top, the burrow openings were dug under the edge of a stone (*Berrill and Stewart*, 1973). In the field, juveniles have been found living in burrows in peat reefs (*Able et al.*, 1988) and were observed digging burrows beneath and between cobbles (*Cobb et al.*, 1983). *Barshaw et al.* (1994) reported juveniles burrowing in peat, mud, and cobble to avoid predation in a laboratory situation. This tendency to burrow suggests that a preference would exist for cohesive, stable mud over sand. If juvenile lobsters preferentially select between mud, sand, and other small grain-size substrates, then determining the strength of this preference is crucial in the development of lobster habitat maps in areas where the seafloor is majority small grain-size.

A common method for mapping the substrate distribution of the seafloor is through acoustics (*Brown et al.*, 2011; *Anderson et al.*, 2008; *ICES*, 2017). Echosounders, both single beam and multibeam, send high frequency sound through the water column to reflect off the seafloor and then record the echo, the characteristics of which can be related to different substrates through acoustic ground discrimination systems (AGDS). Single beam systems are less costly and produce data that are easier to interpret than multi beam systems (*Murphy and Jenkins*, 2010), making them more accessible for small research enterprises. The echo of a single beam AGDS (SB-AGDS) carries characteristics that are indicative of the hardness and roughness of the substrate from which it reflected, allowing for the discrimination between different substrates based on those qualities. Mud and sand are similarly smooth substrates, which is the feature that makes them potentially unappealing to juvenile lobsters, but they are distinct in their levels of hardness; sand is much harder than mud and this difference can be detected by SB-AGDSs, including the one used in this study (*BioSonics*, 2015).

Maces Bay, New Brunswick, is an area of the Bay of Fundy that is known to have a low concentration of cobble (*Fader et al.*, 1977) and some of the most productive lobster fisheries in the region. By combining juvenile lobster sampling and the mapping of the substrate in a subsection of the bay, any relationship between lobster presence and substrate type can be used to develop a species distribution map (SDM) for juveniles. SDMs can be interpreted as a predictor of presence, which is a useful spatial data layer for use in MSP. Similarly, any demonstrated relationship between juvenile presence and small grain-size substrate would prove beneficial for future lobster research and management.

3.2 Materials and Methods

Maces Bay is located on the Fundy shore of New Brunswick, approximately 40 km west of Saint John (45°07' N, -66°30' W). The area of interest is the northwest subsection of the large bay, bounded on three sides by land and spanning from New River Beach Provincial Park to the rural community of Pocologan (Figure 3.1). This subsection is approximately 6 km by 2 km with New River Island (0.13 km^2) in the southeastern quadrant and a series of smaller islands and rock spires in the southwest. A small dammed river flows into the northwest corner via a series of culverts under the highway, and two streams enter on either side of New River Beach in the northeast. There is an active net-pen salmon aquaculture site between the small islands and the western bounding peninsula, and the majority of the activity in the bay is from aquaculture and commercial fishing. The tidal range of the bay is approximately 15 m.

Maces Bay was chosen as the location for this study due to the unique opportunity to obtain georeferenced juvenile presence data. Dinning and Rochette of the University of New Brunswick in Saint John (UNBSJ) conducted a study of juvenile lobsters on the soft sediments of Maces Bay and collaborated with us in their efforts.

3.2.1 Data Collection

3.2.1.1 Acoustic Data

The acoustic data were collected using a vertically-oriented 204.8 kHz 8.6° beam angle transducer and BioSonics Inc. MX aquatic habitat echosounder, which is a single-beam system designed to detect, classify, and map substrate in coastal areas. To maintain consistent coverage and minimize noise interference, the fishing boats on which the system was mounted were kept to a speed of 5 kts (2.6 ms^{-1}). The pings were emitted at a

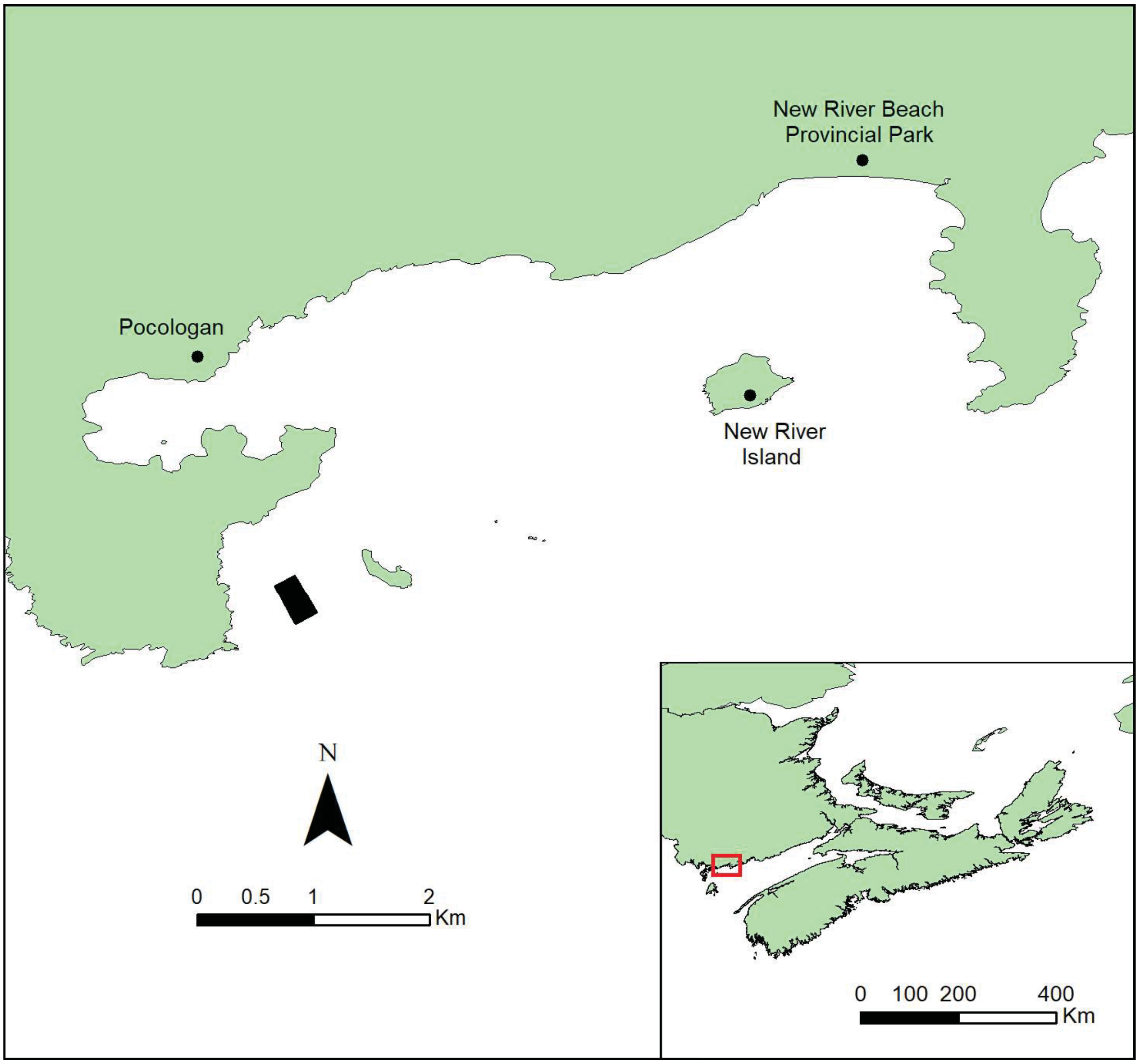

Figure 3.1: Map of the study site, a subsection of Maces Bay. (*Inset*) Location within New Brunswick. The solid black rectangle represents the location of the salmon net pens.

rate of 0.2 Hz, resulting in a sequential spacing of approximately 0.5 m. These were georeferenced to within 2 m with DGPS. The data were collected over the course of several months (October 2015 to July 2016), first in roughly parallel east-west transects separated by approximately 100 m and secondly in roughly parallel north-south transects separated by approximately 50 m (Figure 3.2). The two surveys were conducted independently of each other. The area covered amounted to approximately 8.5 km^2 and 721,861 data points were collected in total.

3.2.1.2 Ground-Truth Video Data

A Seaviewer drop video camera (950 Series, colour version) was used to collect ground-truthing data for both assignment of substrate categories to clustered acoustic pings and verification of the acoustic interpolation. Data collection was conducted jointly with our collaborators at UNBSJ in April and June of 2016 and September of 2017. Videos were taken both on and between the acoustic transects, resulting in 596 video segments georeferenced with a GPS with an accuracy of approximately 2-3 m.

3.2.2 Data Analysis and Substrate Map Creation

3.2.2.1 Acoustic Data

Data were cleaned by removing empty pings and pings with <0.25 m sequential spacing, and correcting the tidal effect on depth using the Lower Low Water Low Tide (LLWLT) datum. Acoustic data were then manipulated into six sets made from combinations of two variables: track spacing (100 m and 50 m) and spatial resolution (2.5 m, 5 m, and 10 m). The spacing and resolution were chosen based on the size of the area to be surveyed, effort limitations, accuracy of the DGPS, distance between sequential pings, and the variance in diameter of the ping footprint with depth. Resolution was derived by binning 5, 10, or 20 pings together, which were referred to as mean pings, or MPs.

BioSonics' software VisualHabitat v2.0.1 was used to analyse the data, applying an unsupervised fuzzy centroid means cluster analysis to group the MPs according to variations in ping echoes that represent the physical characteristics of the substrate. This analysis provided a cluster membership probability for each MP. The primary membership of an MP was the cluster for which they had the highest probability. This information was then processed in a Matlab script to calculate the secondary membership probability for the MPs. This allowed comparison of clusters that shared a large percentage of the same MPs

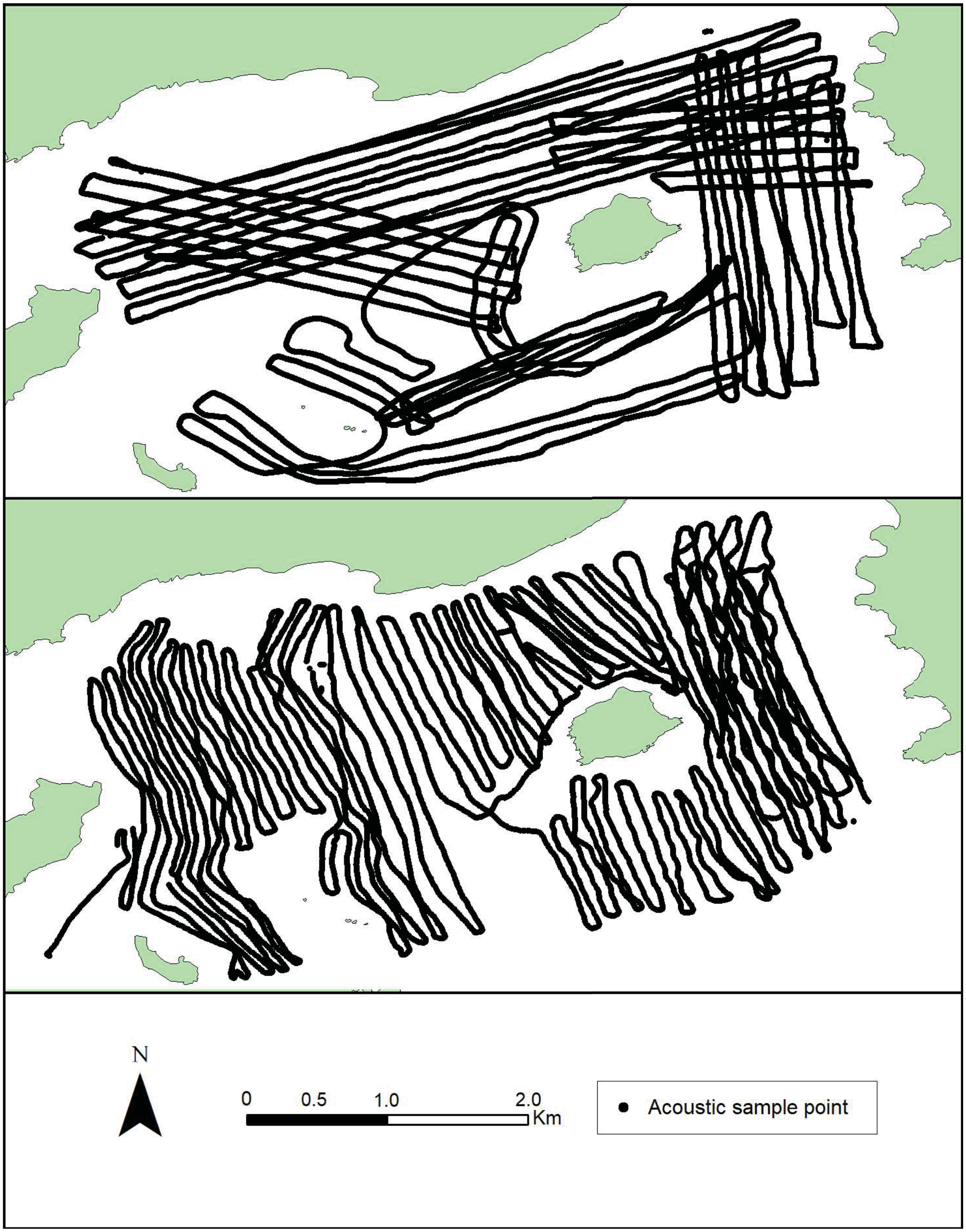

Figure 3.2: Acoustic data survey with tracks are spaced 100 m apart in the top image and 50 m apart in the bottom.

between their primary and secondary memberships and were likely representative of the same substrate. Since this cluster analysis was conducted with a deliberate overestimation of the number of clusters to reduce the chance of conflating substrates, the combining of clusters was an important step in the process and was completed with the help of visual interpretation of the raw echograms. The six sets had either three or four clusters after the combining.

Clusters were assigned substrate categories, the creation of which was based loosely on the work of *Tremblay et al.* (2009). The categories were Rock (RK), Sand (SA), Mud (MD), and Bedrock-Sediment (BS). The first three categories could be differentiated in both the acoustic signal and the underwater video, but BS, which appeared in four of the six data sets, was only differentiated in the acoustic data. Locations with BS MPs appeared as SA or MD in the video, but the acoustic data indicated that the substrate was smooth and rock hard. Therefore, BS represented a thin layer of unknown sediment atop bedrock

3.2.2.2 Ground-Truth Video Data

The 596 georeferenced video segments were each assigned one of the four substrate categories. The BS category was only assigned to a few segments due to its non-visual nature; those segments to which it was applied had some visible indication that the sediment layer was particularly thin. Twenty percent (119) of these video segments were then used to assign substrate categories to the clusters in each of the six data sets. This subsample of video segments was randomly selected from the videos that shared spatial coordinates with acoustic data. Those videos that showed MD or SA but were at the same location as the clusters that inspired the creation of the BS category were assigned BS – the unknown classification of the sediment above the bedrock meant that SA and MD could both be valid visual indicators of the acoustically-defined category.

3.2.2.3 Substrate Maps

The distribution of substrate categories throughout the bay was completed using indicator kriging, where the semivariograms were defined with a spherical model, the automatic detection of parameters, and a smoothing factor of 0.2, with no identified anisotropies. The results of this process were raster maps with each grid cell assigned the substrate category that had the highest probability of occurring in that location. This process was completed for all six data sets using ArcMap v10.5 in the WSG84 datum.

3.2.3 Trawl Data

The substrate maps of the surveyed area were then used to trawl for juvenile lobsters (<40 mm CL) between July and November of 2016. The survey was deliberately biased in favour of small grain-size substrate and trawled approximately a third of the surveyed area with a modified shrimp trawl approximately 1 m wide at an average speed of 1.5 kts. The results of each trawl were recorded as present or absent instead of as abundances. The substrate maps were then reduced to the area the trawls covered with an 80 m buffer.

The "absent" label used for the trawl results was not indicative of a true absence for the purposes of this study due to the difficulty in determining true absence. True absence indicates that the location sampled is unsuitable habitat for the species (*Elith et al.*, 2010). However, the multitude of other possible explanations for absence, including competition, population density maximums, dispersal limitations, and local extinctions, can make that distinction unclear (*Pulliam*, 2000). Since I was not confident that these absences were true absences, and the software used to create the end models only required presence data, the absences were instead considered pseudo-absences.

3.2.4 Validation and Uncertainty

Two readily quantifiable sources of uncertainty were then examined as a way of selecting which of the six maps to use in further analysis: the classification certainty, which is a measure of the agreement between the substrate interpolation and the ground-truth data; and modelling certainty, which is the level of certainty in the prediction created by the interpolation.

The classification certainty was resolved through error matrices that allowed the direct comparison of the unused 80% of the ground-truth videos with the substrate interpolation (Table 3.1). However, since the area of interest had reduced substantially from the full survey to the trawled area, only those video segments that were both within the smaller bounds and a part of the remaining 80% were used (between 63 and 96 video segments, depending on the resolution and track spacing). The agreements and disagreements between the ground-truth data set and the interpolation were revealed in error matrices which accounted for each of the 16 possible outcomes. These 16 outcomes can be characterised as four types relative to the substrate category under consideration. For the MD category: true positives (TP), where the two data sets agreed on the presence of MD, were recorded as V_1A_1; false positives (FP) or errors of commission, where the presence of MD was falsely predicted by the interpolation, were comprised of V_2A_1, V_3A_1, and V_4A_1; false negatives (TN) or errors of omission, where the absence of MD was falsely predicted, were in V_1A_2, V_1A_3, and V_1A_4; and true negatives (TN), where the data sets agreed on the absence of MD, were the remaining nine outcomes. These results allowed for the calculation of further error statistics (Table 3.2).

In the creation of the substrate maps, the substrate category with the highest probability was selected for each cell. That probability represented the certainty involved in that selection and could therefore be considered as the modelling certainty. This was represented as a map layer.

From these two processes, the data set selected as the most accurate was the one composed of pings collected on tracks spaced 50 m apart and with a spatial resolution of 10 m. This map was then subjected to further analyses.

3.2.5 MaxEnt

MaxEnt environmental niche modelling software (*Phillips et al.*, 2006) was used to develop SDMs from the data. It requires both environmental variable rasters, represented by the substrate map and the depth raster (created from the acoustic data through interpolation), and georeferenced presence-only species data, represented by the midpoints of the trawls with recorded juvenile presence. Given the intent to examine the preferences of juvenile lobsters on sand versus mud, the RK category was removed from the substrate map before analysis; BS was not found within the reduced (trawl-sized) map and was therefore not considered.

This SDM analysis was conducted twice, once with the substrate data as selected above, referred to as the Absolute model, and once with an additional substrate category introduced in order to explore the effect of heterogeneity on the species occurrence, referred to as the Heterogeneous model. In the Heterogeneous model, those areas that had a modelling certainty of equal to or less than 30% were relabelled as "Heterogeneous" or HG. Areas that had a modelling certainty higher than 30% kept their original substrate category classification. This created a new substrate map composed of SA, MD, and HG that was then processed in MaxEnt.

Table 3.1: A) The assignation of each substrate category in the two data sets to a placeholder label. B) An error matrix representing the sixteen possible outcomes of the comparison between the acoustic interpolation and the video segments.

A.

Video	Acoustic
V_1 = MD	A_1 = MD
V_2 = SA	A_2 = SA
V_3 = BS	A_3 = BS
V_4 = RK	A_4 = RK

B.

Acoustic Interpolation (Prediction)	Video (Ground-truth)			
	V_1	V_2	V_3	V_4
A_1	V_1A_1	V_2A_1	V_3A_1	V_4A_1
A_2	V_1A_2	V_2A_2	V_3A_2	V_4A_2
A_3	V_1A_3	V_2A_3	V_3A_3	V_4A_3
A_4	V_1A_4	V_2A_4	V_3A_4	V_4A_4

Table 3.2: Descriptions and equations of the statistics derived from the error matrices (adapted from *Barrell et al.* (2015))

Statistic	Equation	Description
Overall Accuracy	$\frac{TP+TN}{n}$	Proportion of all predictions that were correct
Sensitivity	$\frac{TP}{TP+FN}$	Proportion of correctly predicted presences
Specificity	$\frac{TN}{TN+FP}$	Proportion of correctly predicted absences
Positive Predictive Value (PPV)	$\frac{TP}{TP+FP}$	Proportion of positive predictions that are TP
Negative Prediction Value (NPV)	$\frac{TN}{TN+FN}$	Proportion of negative predictions that are TN

3.3 Results

3.3.1 Substrate Maps

The map with the 50 m track spacing and 10 m resolution was selected for the full analysis and had 15,502 MPs, which broke down into 6249 categorised as MD (40.3%), 4085 as SA (26.4%), 1847 as BS (11.9%), and 3321 as RK (21.4%) (Figure 3.3). The categorised MPs display clear groupings: RK to the south and east of New River Island and around the smaller islands in the west; MD in the northeast and northwest corners; SA in the north centre; and BS in the far west. The margins of the MD and SA areas in the northern half are highly intermixed. The RK area in the southeast has a substantial amount of BS within it, particularly on the southern shore of the island; there is also a distinct patch of SA within the RK.

The interpolated map of the entire surveyed area displayed many of the same substrate specific areas highlighted above (Figure 3.4). The patches of BS and SA within the RK area in the southwest came through the interpolation, implying that the differences in substrate are true and not anomalies. The gradation between the MD and SA areas was turned into definite boundaries, resulting in the somewhat irregular boundary between the northeast MD and the north centre SA. Similarly, there was a tiny patch of SA in the northeast MD section, which is likely the result of the intermixed MD and SA MPs that existed there before the interpolation.

3.3.2 Trawls

One hundred and fifty five trawls ranging in size from 13 m to 777 m were successfully completed within the surveyed area (Figure 3.4), 89 of which had juvenile lobster presence (56.7%). The area covered by the trawls was mostly predicted to be MD and SA, with the exception of some RK areas on the edges (Figure 3.5). The trawls frequently crossed over the boundaries between MD and SA, sometimes more than once on the same trawl.

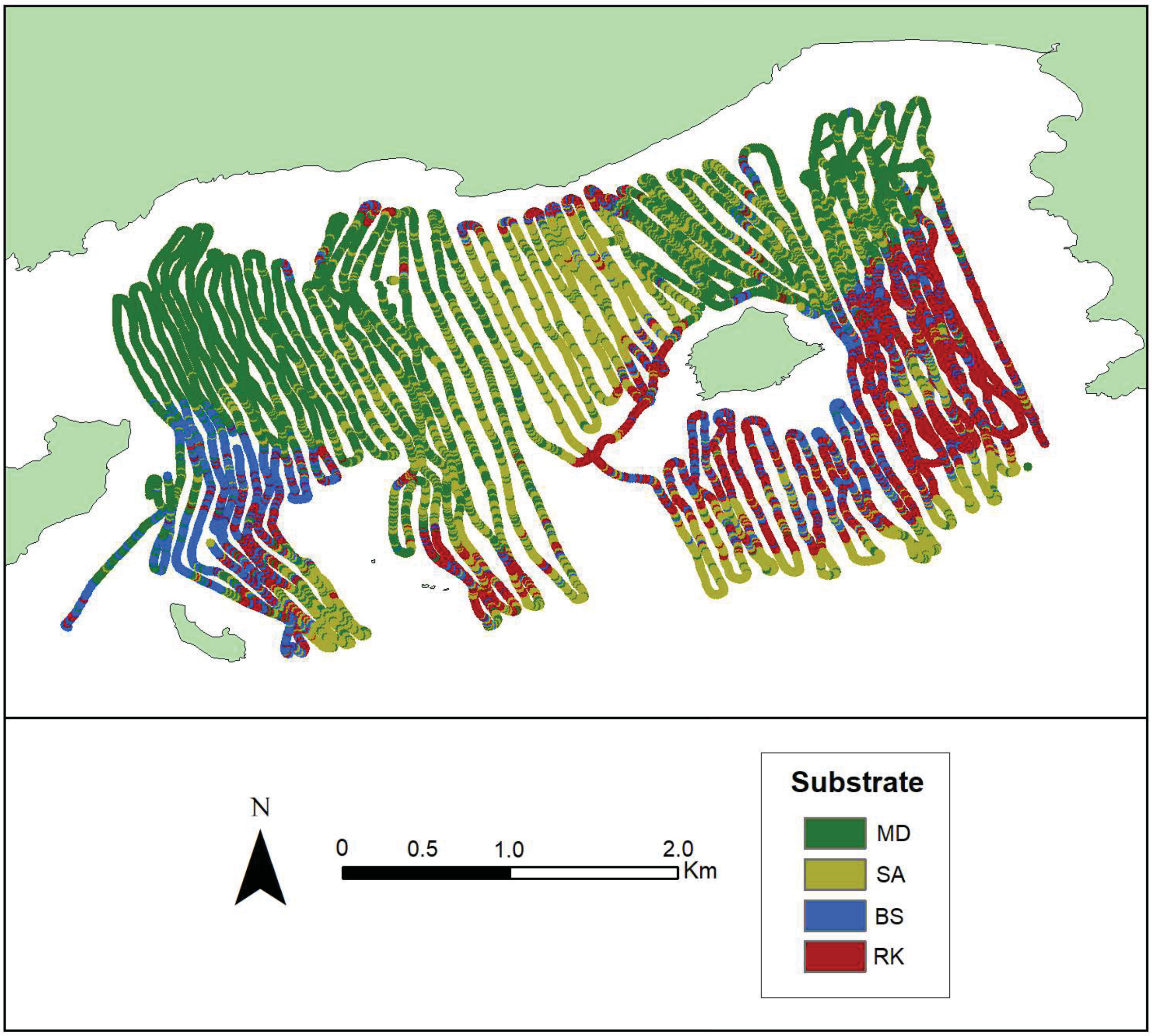

Figure 3.3: Map of acoustic MPs categorised into substrate categories. The tracks in this map are spaced 50 m apart, and the spatial resolution is 10 m.

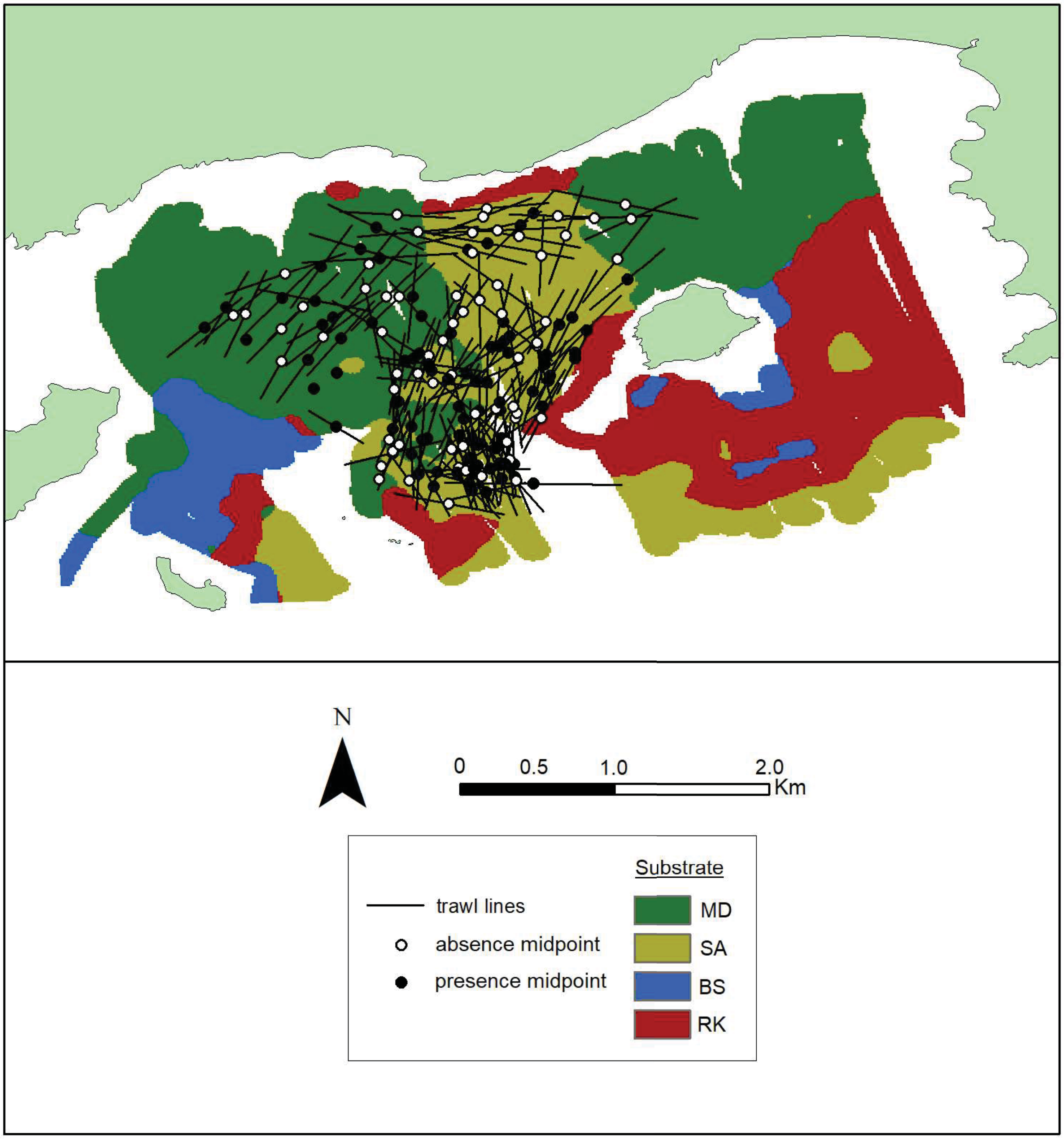

Figure 3.4: Substrate map of the full surveyed area, made from tracks spaced 50 m apart and with a spatial resolution of 10 m. The midpoints of the trawl lines denote whether the trawl found presence of juvenile (<40mm CL) lobsters.

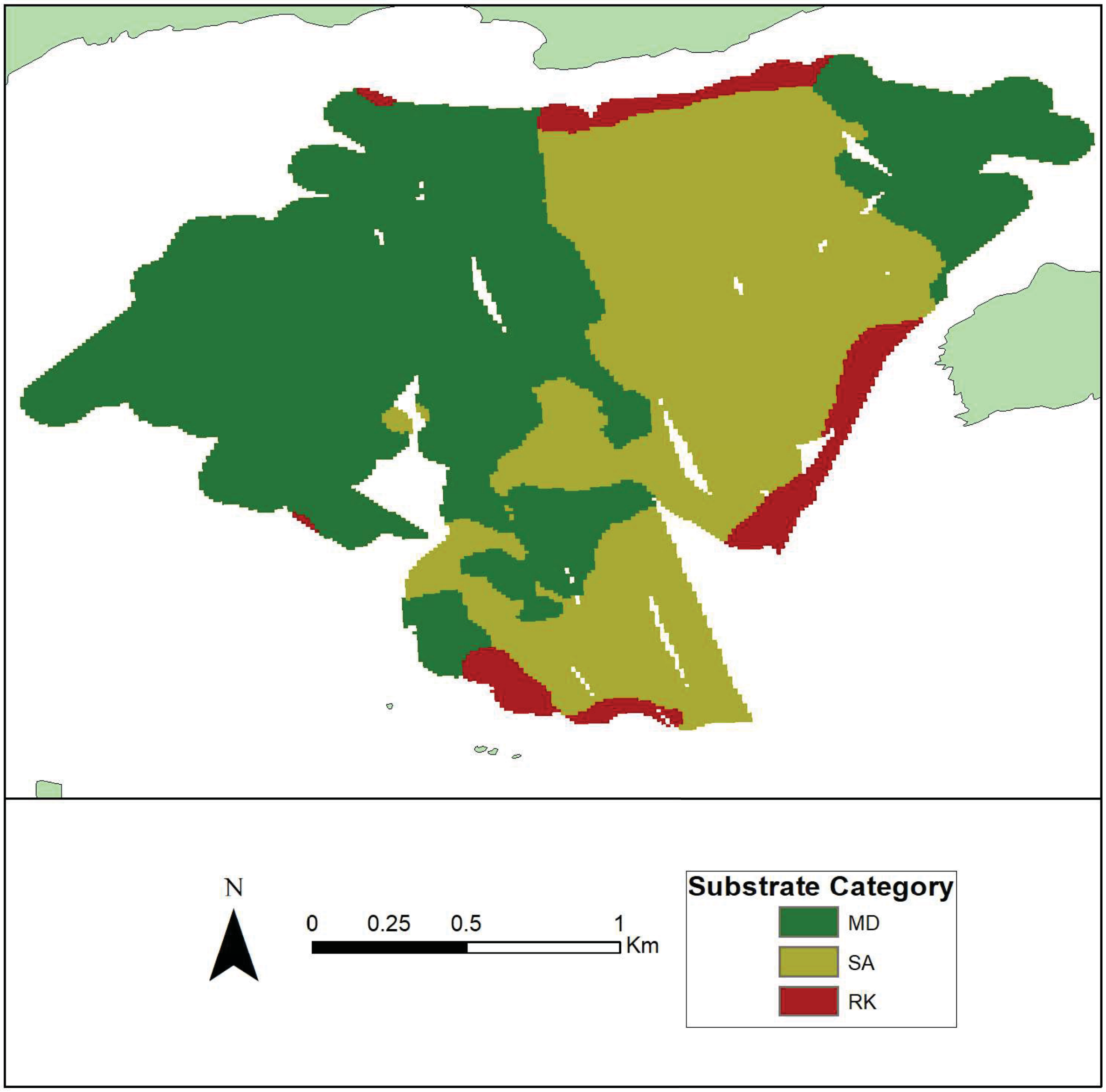

Figure 3.5: Predicted substrate distribution beneath the trawl lines, with an 80 m buffer.

3.3.3 Validation and Uncertainty

The classification error matrix and statistics (Table 3.3, 3.4) show that the RK and MD categories both had high sensitivities (0.9444 and 0.8947, respectively), indicating that they were rarely incorrectly predicted by the interpolation. Similarly, high PPVs (0.7727 and 0.9189) meant that prediction of both RK and MD were rarely incorrect. This model was clearly able to accurately predict for MD and RK distribution but it faltered with SA, which was incorrectly predicted as RK 36.4% of the time and as MD 18.2% of the time, resulting in a low sensitivity value of 0.4545. These values may have been partially due to the relatively small number of SA videos, of which there were only 11; one incorrect prediction has more weight when the total number of predictions is low. Of the eight SA predictions made, three were actually MD video segments, meaning SA's PPV is also significantly lower (0.625) than either RK or MD. This confusion regarding SA is not unexpected, as sand possesses physical characteristics similar to both rock (hard) and mud (smooth). However, the overall accuracies for all the categories were more affected by the number of TNs than the TPs, meaning that SA had a relatively high overall accuracy of 0.8657.

This data set with 50 m spacing and 10 m resolution was best at correctly predicting all three categories (BS was not present in the map), especially SA, and was best at differentiating between MD and SA. It also had the lowest number of classification errors when compared to the other five data sets (22 versus 36–56). The modelling certainty was highest of the six (Figure 3.6), with certainties of up to 80% in the SA area northwest of the island and up to 100% in the MD in the far western section. The boundaries are of low certainty, particularly in the southern section.

Table 3.3: Error matrix representing the comparison between the acoustic interpolation predictions and the underwater video.

Acoustic Interpolation	Video (Ground-truth)			
(Prediction)	MD	SA	BS	RK
MD	34	2	0	1
SA	3	5	0	0
BS	0	0	0	0
RK	1	4	0	17

Table 3.4: The error statistics derived from the agreements and errors revealed in the error matrix.

	Overall Accuracy	Sensitivity	Specificity	PPV	NPV
MD	0.8955	0.8947	0.8966	0.9189	0.8667
SA	0.8657	0.4546	0.9464	0.6250	0.8983
BS	1	–	1	–	1
RK	0.9105	0.9444	0.8980	0.7727	0.9778

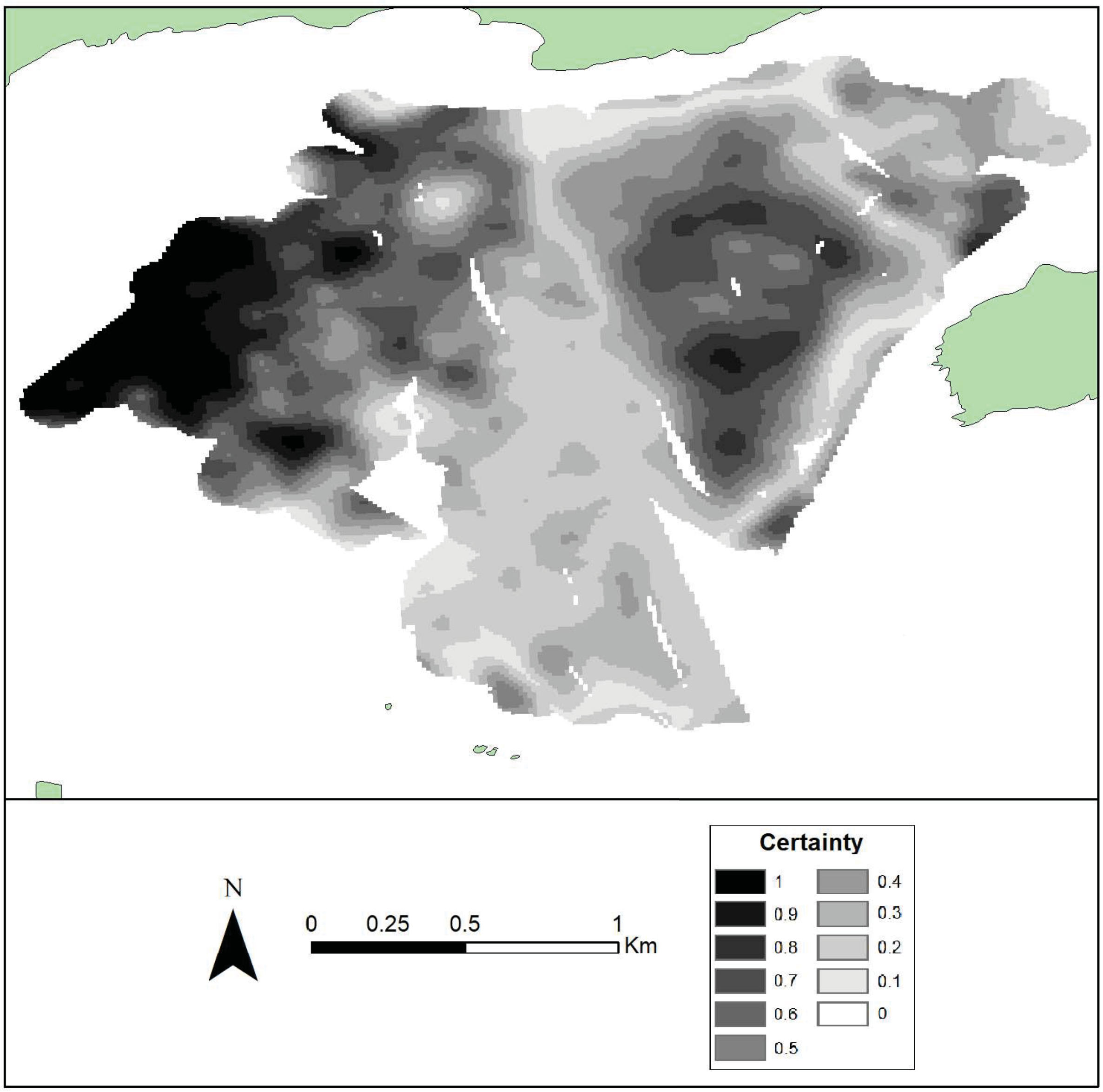

Figure 3.6: Modelling certainty maps derived from the probability of the interpolated data.

3.3.4 MaxEnt

I used the raw output from MaxEnt, which calculates the species response in units of "relative occurrence rate" or ROR. The values are set relative to other cells and therefore provide a qualitative measurement of the suitability of the surveyed area.

The SDM of the Absolute model displays a clear trend of higher ROR values in the southern part of the map that fade into lower values moving further north, with the extreme low values in the northeast corner above New River Island (Figure 3.7).

There is no visible distinction made between the two substrate categories, a finding which is supported by the species response graph (Figure 3.8). This graph clearly illustrates the confounding nature that depth, which ranges from -3.4 m to -21.2 m, has on the effect of substrate. The difference between the average responses to SA ($1.44\text{x}10^{-4}$) and MD ($1.48\text{x}10^{-4}$) was very small when depth was included as a variable, but the change in response that occurred when the depth variable was excluded (SA: $1.29\text{x}10^{-4}$; MD: $7.63\text{x}10^{-5}$) is indicative of depth playing a larger role in species response over all. This is confirmed by the MaxEnt statistic estimating depth's contribution to the model at 100%.

The higher ROR values corresponded with deeper depths (Figure 3.7) but also with lower modelling certainty (Figure 3.6), which implies heterogeneity. It was this relationship which led to the development of the Heterogeneous model, where the heterogeneous substrate category, HG, was added to replace all the areas with a modelling certainty of 30% or less (Figure 3.9).

The SDM for the Heterogeneous model (Figure 3.11) displayed an almost identical pattern of RORs as the Absolute model, with higher values in the south and lower values in the north. Again, there was no visible distinction between the SA and MD areas. This was again supported by the species response graph, where there was almost no differentiation between the three categories when depth was left in the calculation, with HG and MD both at $1.46\text{x}10^{-4}$ and SA at $1.53\text{x}10^{-5}$. When depth was removed, HG prompted the highest response at $1.25\text{x}10^{-4}$, closely followed by SA at $1.17\text{x}10^{-4}$, and then MD at $5.36\text{x}10^{-5}$. This significant change in the response pattern following the removal of the depth variable implies that the depth is once again the stronger variable in eliciting species response, which the MaxEnt contribution statistic confirms by estimating 99.9% contribution by depth to the model.

The receiver operating characteristic (ROC) curves, which measure how well the model performs compared to the results of a theoretical model based on random distribution predictions, are almost identical for both models (Figure 3.10). The area under the curve (AUC) for the Absolute model is 0.714 and for the Heterogeneous model is 0.713, which are both better values than the random model (AUC = 0.5) but not exceptionally so.

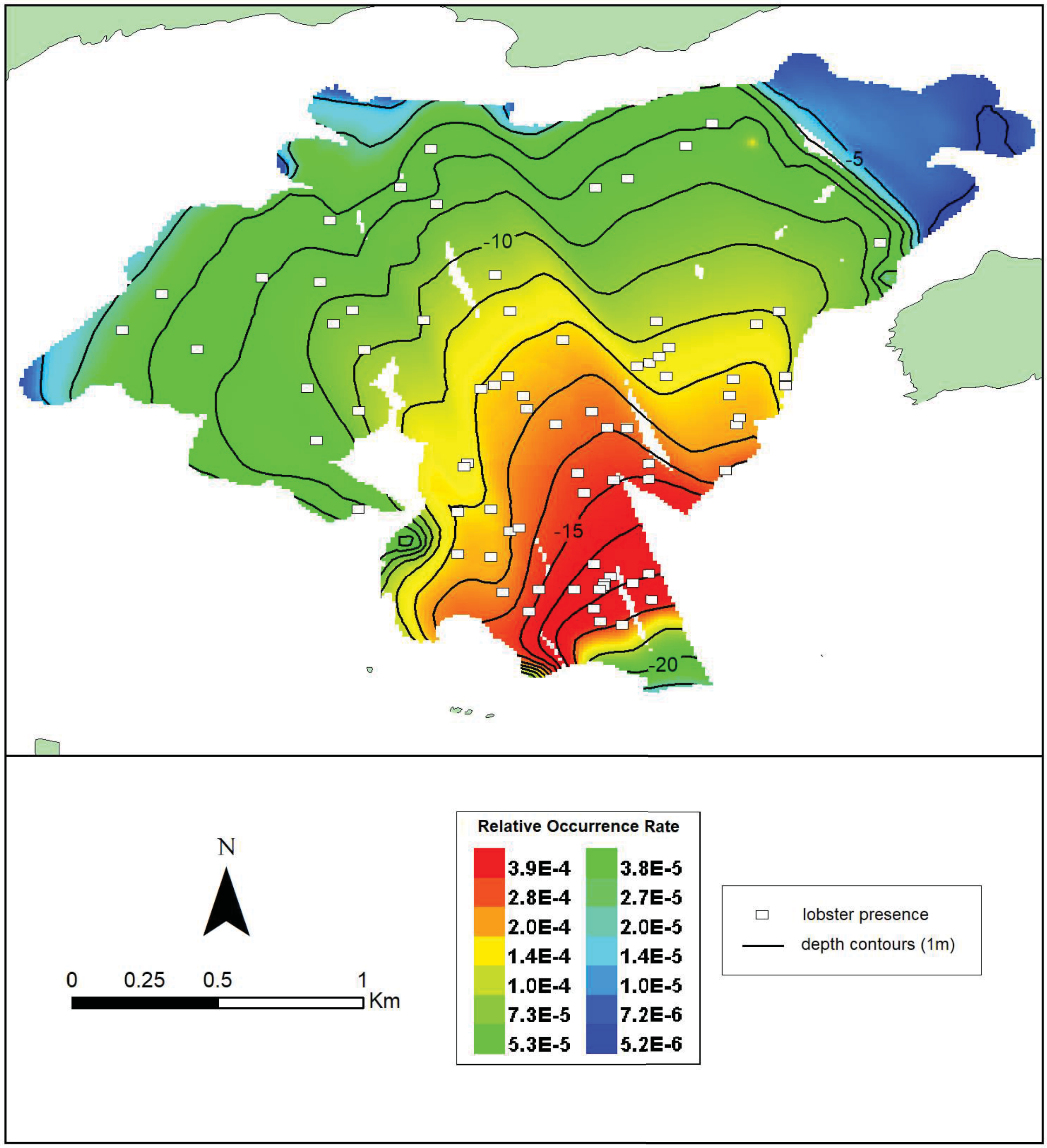

Figure 3.7: SDM of the Absolute model. The white rectangles indicate the midpoints of the lobster trawls and the depth contours are 1 m spaced.

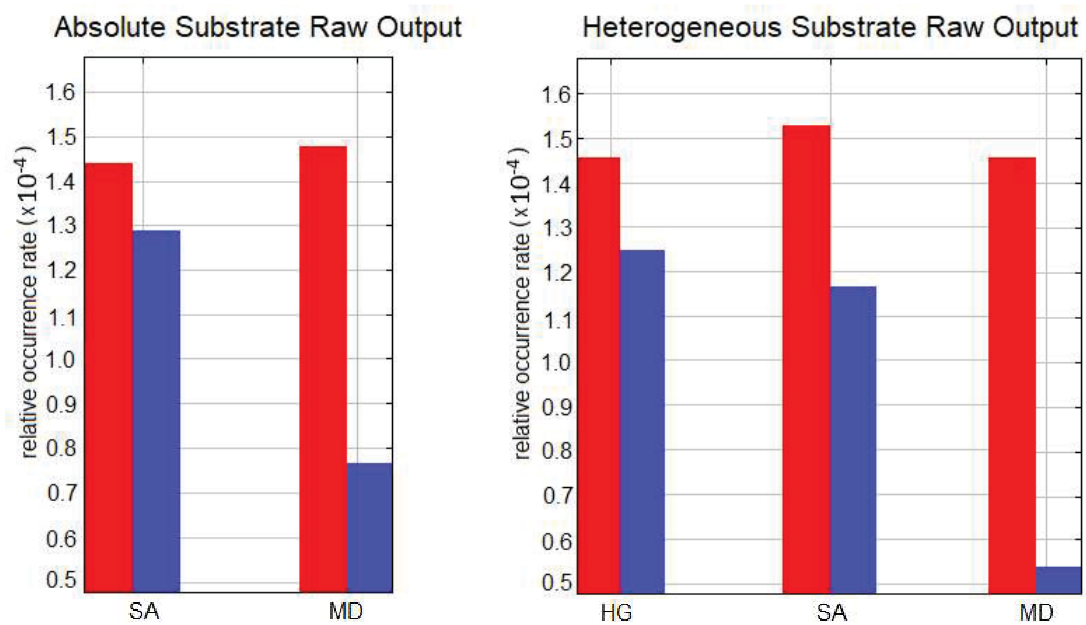

Figure 3.8: Graphs of species response to substrate categories for both the Absolute and Heterogeneous models. The red is the marginal response to the substrate variable with the depth variable included; blue is the response to the substrate variable without the depth variable.

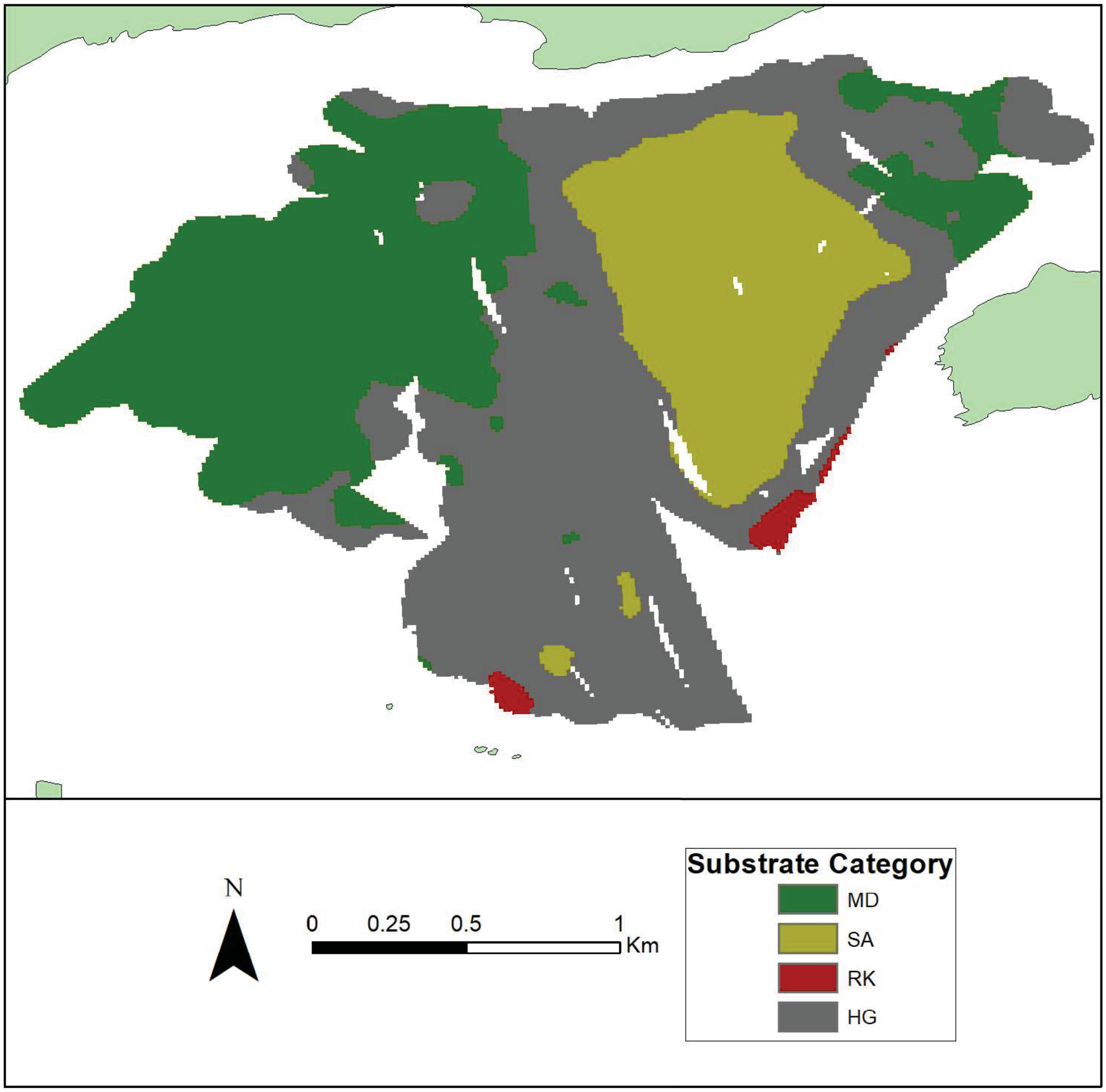

Figure 3.9: Figure 3.9: Substrate map of the Heterogeneous model, where HG represents heterogeneous substrates defined as having a modelling certainty of 30% or less. The other substrates are defined in the same manner as the Absolute model.

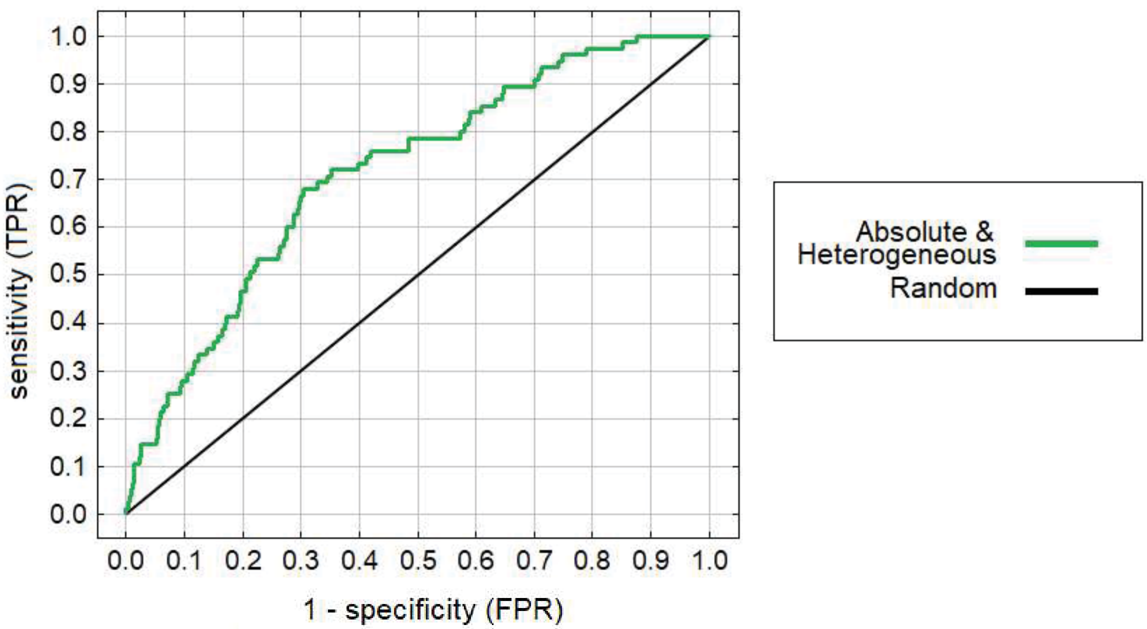

Figure 3.10: Receiver operating characteristics (ROC) curve for both the absolute (AUC = 0.714) and heterogeneous (AUC = 0.713) models. The curves for the two models are similar and the differences cannot be accurately displayed, so the single curve is representative of both models.

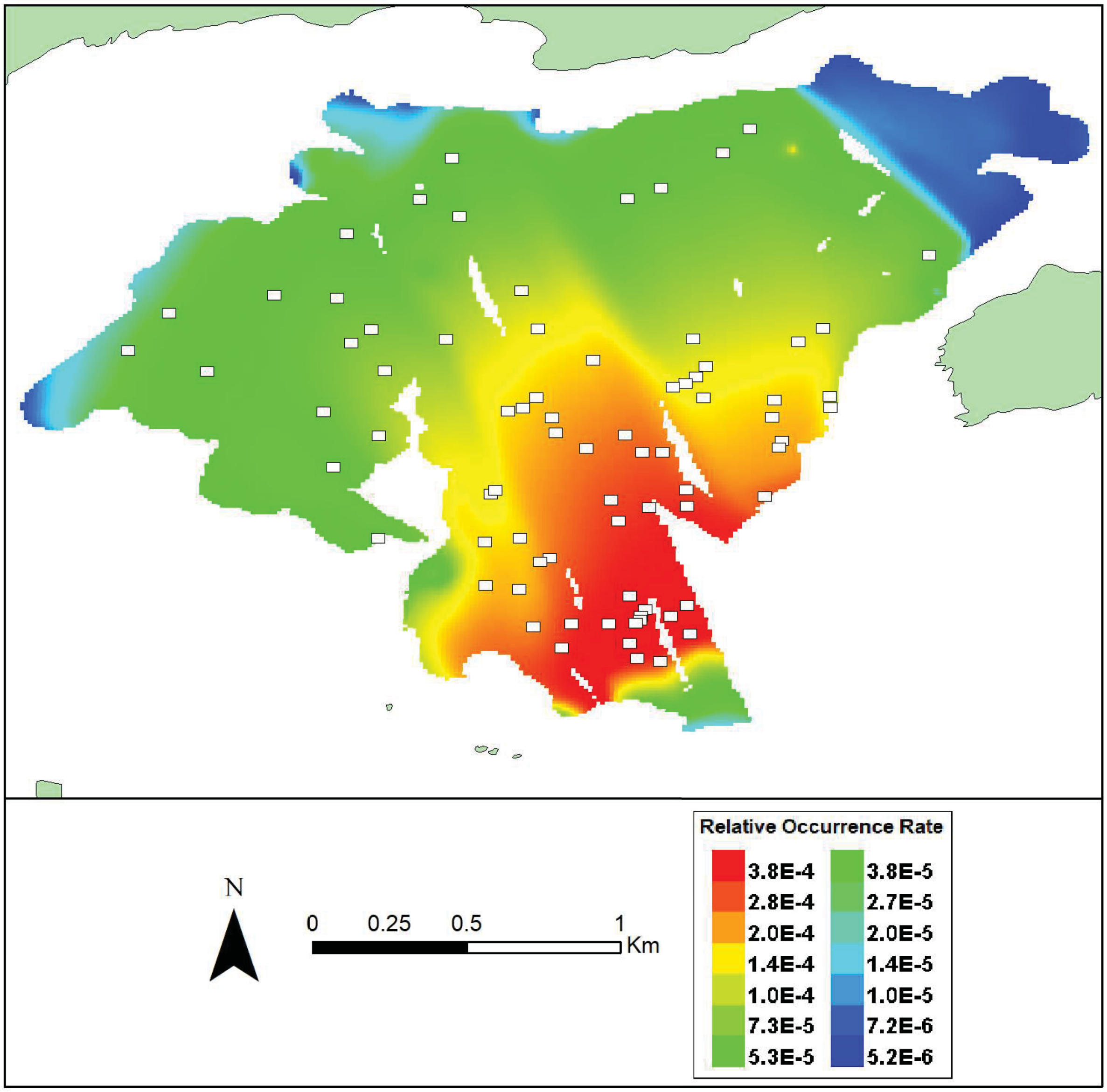

Figure 3.11: SDM of the Heterogeneous model. The white rectangles indicate the midpoints of the lobster trawls.

3.4 Discussion

This study mapped the suitability of fine grain substrates for juvenile American lobsters in Maces Bay, NB, using a SB-AGDS, video ground-truthing, and trawled lobster presence data. The substrate classifications had high certainty (Tables 3.3, 3.4) and were satisfactory for the model, and the georeferenced presence data were collected from the same location as the acoustic data, so the credibility of the map is sound.

The consistent spatial presence of juveniles in areas classified as fine grain sediments challenges the understanding of juvenile lobster settlement, previously thought to occur largely on cobble. This will consequently affect future methods of habitat mapping for juveniles. Similarly, the absence of clearly demonstrated differences in suitability between mud and sand introduces questions about the nature of substrate preference in juvenile American lobsters.

3.4.1 Substrate Models

The lack of differences between the Absolute and Heterogeneous models (Figure 3.10) is indicative of the relative strength of depth as the main predictive variable in comparison to the substrate categories selected. The HG substrate category spans almost the entire depth range of the surveyed area but there are more presence points within the deeper southern section than in the shallower northern sections. The HG category could be separated into two categories, one based on heterogeneity with cobble or gravel patches in the southern section (far south RK patches in Figure 3.5), and the other based on heterogeneity including only the gradient between MD and SA. This separation of HG may have elucidated a response stronger than that of depth, but it would likely have been the result of cobble and gravel being preferred habitat (*Pottle and Elner*, 1982) and not due to any heterogeneity of MD and SA. This indicates that the depth effect occurs as a confounding influence of cobble in the deepest part of the mapped area.

3.4.2 Substrate Importance

Considering the suspected preference juvenile lobsters have between small grain-size sediments (*Botero and Atema*, 1982), the lack of preference demonstrated in the habitat suitability maps could be explained through two different viewpoints: (1) the difference between the substrates is not substantial from the perspective of the lobsters and therefore there is no response difference (false distinction); and (2) there is a response difference but it was not detected through this method (technological impairment).

3.4.2.1 False Distinction

While juveniles in low cobble areas may be unable to access cobble due to intense competition (*Steneck*, 2006; *Tang et al.*, 2015; *Wahle and Steneck*, 1991), it is unlikely that the fine sediments they must inhabit are of equally low quality. There is a dearth of records of juvenile lobsters found on clean sand, both in the lab and in the field (*Botero and Atema*, 1982; *Cobb et al.*, 1983; *Able et al.*, 1988; *Barshaw et al.*, 1994), and when compared to the limited but consistent literature showing juveniles digging burrows in mud in the absence of cobble (*Berrill and Stewart*, 1973; *Barshaw and Bryant-Rich*, 1988; *Cobb et al.*, 1983), a preference between mud and sand is likely to exist. Therefore the lack of demonstrated preference in the SDMs is probably due to a lack of a functional difference in the surveyed substrates from the perspective of the juvenile lobsters.

The substrate in the northwestern section of the Bay of Fundy that includes Maces Bay consists largely of mud and sandy mud (*Fader et al.*, 1977; *Shaw et al.*, 2012), with limited patches of cobble, boulders, and sand. The SA category detected in this study could be clean sand or it could be sandy mud. If SA has a high mud content, there may be no discernible difference in the ability of juvenile lobsters to create burrows in SA as compared to mud. Therefore, SA may not be different enough from true mud to affect juvenile habitat selection despite it being a valid category and distinct from better sorted silt clay in the acoustic sediment classification system. Since it is not known at what point in the sand-to-mud gradient juvenile lobsters begin to reject the sandier sediment and show preference for the muddier one, it is possible that my delineation of categories is mismatched with that of the preferences of lobsters. This mismatch could result in the juvenile lobsters seemingly showing no preference between mud and sand because both MD and SA are cohesive enough sediments to allow burrows.

While I could find no published literature on the explicit preferences of grain-size of the American lobster, relevant studies have been conducted for the European lobster (*Homarus gammarus*). When presented with substrates that ranged from unsieved mud (<0.06 mm grain size) through various coarseness levels of sand up to small rocks (7–20 mm), stage VII *H. gammarus* selected rocks 92% of the time and mud 100% of the time (*Howard and Bennett*, 1979). Sands experienced between zero and 33% selection, and the smallest sand category of 0.06–0.25 mm was selected only 25% of the time. This indicates a similar aversion to sand as postulated to occur for *H. americanus*, and the extensive similarities between *H. americanus* and *H. gammarus* (*Cobb and Wahle*, 1993) suggest that the grain size preferences may also be shared. However, experimentation into detailed grain size preferences of *H. americanus* would be valuable to establish more certainty.

3.4.2.2 Technological Impairment

Trawling is the only practical method for surveying juvenile lobsters over large areas. While converting line data into point data based on the midpoint location is a common method of handling trawls (*Turner*, 2014; *Rooper et al.*, 2014; *Petitgas et al.*, 2003), it creates uncertainty by averaging any location effects into a single point. In this study, location is inherently tied to the environmental variable being examined (ie. substrate), so the loss of precision resulting from the averaging is an unknown source of variance. In particular, trawls that cross from one substrate to another may falsely attribute presence to the incorrect substrate due to the location of the midpoint. Additionally, while MaxEnt requires presence-only data, abundance counts per trawl would have allowed a comparison to be made between the number of lobsters caught and the percentage of the trawl spent over each substrate, possibly leading to the discovery of a pattern. Abundance data from the trawls were not available for this study.

The descriptive resolution of the drop camera may explain the lack of substrate-driven species response. "Descriptive resolution" is a term that describes the level of habitat detail a sensor is able to reveal (*Green et al.*, 1996). The drop camera used in ground-truthing was only able to reveal broad descriptive resolutions in the fine end of the substrate spectrum, limiting the categories to sand and mud based on the respective presence and absence of ripples in the sediment (*Whitlatch*, 1977) and the relative settlement rates of resuspended sediment. Despite the acoustic signal displaying an apparent gradient between MD and SA, an attempt to further resolve mud and sand was difficult because they could not be validated through video ground-truthing. A regime of grab samples would have been useful in handling this problem, but was not an accessible method at the time of data collection.

3.4.3 Effects of Fine Sediment Habitat Use on Lobster Population

The increased use of less preferred substrates by juveniles in the past couple decades is likely due to population growth (based on landings: (*DFO*, 1995, 2005, 2015)) resulting in increasing competition for cobble (*Tang et al.*, 2015; *Steneck*, 2006; *Wahle and Steneck*, 1991). This change in habitat could potentially have effects on the health and structure of the lobster population. *Tang et al.* (2015) examined the size-at-age and body condition of juveniles collected on mud and cobble in Maces Bay and found that individuals collected on mud were the same size as older individuals collected on cobble, which they hypothesised may be due to differences in diet between the two habitats. The direction of this size disparity aligns with the growth inhibition effect of competition reported by *Cobb and Tamm* (1974) and *Nelson et al.* (1980), where lobsters in crowded environments grew at a slower rate than those in less densely populated environments.

Additionally, increased habitation of mud plains could lead to other effects on the population. Muddy sediments are susceptible to hypoxia, though burrows are typically better oxygenated than surrounding sediments due to ventilation (*Aller*, 1988). Most burrowing decapod species demonstrate increased tolerance and adaptation to low oxygen environments (*Bridges and Brand*, 1980). This includes the Norwegian lobster (*Nephrops norvegicus*), a member of the same family as *H. americanus* that lives almost exclusively in burrows in deep water mud, and exhibits physiological adaptations that allow it survive both acute and chronic hypoxia events (*Hagerman and Uglow*, 1985; *Baden and Neil*, 2003). Juvenile *N. norvegicus* are more susceptible to hypoxia than adults (*Eriksson and Baden*, 1997). *H. americanus* does have the ability to survive acute moderate hypoxia (*McMahon and Wilkens*, 1971, 1975), but little is know about the effects this may have on long term or population health. Given the relative rarity of *H. americanus* burrowing into mud, it is possible that it has not evolved to manage hypoxia as well as *N. norvegicus*, and *H. americanus* juveniles inhabiting mud burrows may suffer negative effects of low oxygen. Considering the potential importance of soft sediments to juvenile American lobsters, it is surprising that so little is known of their animal-sediment relationships. Laboratory experiments would be invaluable to better understand the relative significance of mud as a substrate for juvenile American lobsters.

3.4.4 Representative Subsample and Future Directions

Given the other substrate maps of the region (*Shaw et al.*, 2012; *Fader et al.*, 1977), it is likely that the substrate map developed in this study is a broadly representative subsample of the larger areas of Maces Bay and the northwestern Bay of Fundy. That is, there are expansive plains of sediment of various levels of grain size, interspersed with the occasional cobble, gravel, or boulder patch. Similarly, the juvenile lobster presence is likely to be representative of juvenile populations in the areas where there is competition for small patches of ideal substrate; the trawl catch rate of juvenile lobsters would likely be far smaller in regions with large amounts of cobble and gravel patches than regions with only smaller grain-size sediment. Both of these conclusions give support to the suitability of generalising the SDMs to a larger area, with the understood caveat that depth is not necessarily a true predictor of suitability.

Future attempts to spatially resolve the relative habitat suitability of sand and mud for juvenile American lobsters in low cobble areas should include a stronger understanding of both the details of grain-size preference among juveniles, and the further resolution of the substrate signal to match. This increased substrate resolution could be obtained through a grab sampling regime in combination with paired echosounders, one 50 kHz and one 200 kHz to increased the range of detectable substrate characteristics (*Freitas et al.*, 2008).

3.5 Conclusions

This study has shown that the use of SB-AGDS was effective in creating fine scale juvenile American lobster habitat suitability maps of small grain-size substrates. However, the lack of species response due to substrate suggests that there is widespread suitable habitat available, largely because there is sufficient mud content in the sediments for burrow formation and stability.

This study can also be taken as evidence that juvenile lobsters inhabit less preferable substrate types, giving support to the theory that sedimentary substrates can act as nurseries in the absence of adequate pre-formed shelter. Due to the importance of soft substrates in this regard, management of lobster stocks must include consideration of a wider range of benthic habitats as potential nursery grounds.

CHAPTER 4

CONCLUSION

The lack of species habitat map layers for use in planning purposes is a major stumbling block in structuring the local use of ocean space. Marine spatial planning requires spatial data that has frequently not been collected or is only available at inappropriate scales, hampering its effectiveness. With MSP potentially providing a partial solution to the social conflicts existing locally between the net-pen salmon aquaculture industry and the lobster fishery, the integrity of the data included in the planning is paramount. This thesis explored the applicability, limitations, and accuracy of using SB-AGDS, underwater video, and two methods of presence sampling to create SDMs for the future use in MSP.

Chapter 2 showcased a viable method of producing adult lobster SDMs for MSP programs through the use of SB-AGDS, underwater video, and georeferenced lobster trap buoys. The results show that in environments where the acoustic signals of the substrates are of suitably high contrast, substrate maps can be created. The presence sampling of the adult lobster population through lobster trap buoys was a convenient and useful method, despite the existence of important cautionary considerations in its use. Depth appeared as a confounding factor in the SDM, but the overall distribution structure of the map was valid. This chapter also explored the effect of spatial resolution and track spacing on the interpretation of the data. The results dramatically expressed the importance of spatial scale analysis, showing meaningful differences in both substrate distribution and species response depending on the spatial scale used. The limitations of the spatial methods in highlighting areas of patchy substrate was addressed, as was the potential conflicts in substrate categorisation.

Chapter 3 demonstrated the effectiveness of the joint SB-AGDS and video method for juvenile American lobsters living on fine substrates, with the notable difference in the method of collecting lobster presence data. These results indicate that the juveniles are present on fine substrates, which is generally counter to what has been reported in the literature, and have no apparent preference between the two fine substrates resolved by the SB-AGDS. This suggests that a more thorough understanding of what constitutes suitable habitat for juveniles is required to accurately model their distribution.

As a whole, this document demonstrates that creating SDMs can be a complex yet accessible process if one pays close attention to the changing minutiae of the environmental variables, the species and their life stage, the various resolutions, and the effect of confounding factors. The methods, techniques, and conclusions derived in this thesis can and should be considered specifically by those developing MSP programs with benthic species in the Canadian Maritimes and nearby regions, and more generally elsewhere.

APPENDIX A

MAXENT

Species distribution models (SDMs) are a type of habitat map which describe the suitability of an environment for a species through mapped environmental conditions. This is done by layering measured environmental variables with presence data, and replicating the patterns to predict presence in unsampled areas (Figure A.1).

MaxEnt is species distribution and environmental niche modelling software developed by *Phillips et al.* (2006) that produces SDMs. It uses the principle of maximum entropy to estimate species distribution across geographic space . In doing this, it expands the realised niche of the species, which is represented by the presence data, into a prediction of its fundamental niche (*Phillips et al.*, 2006).

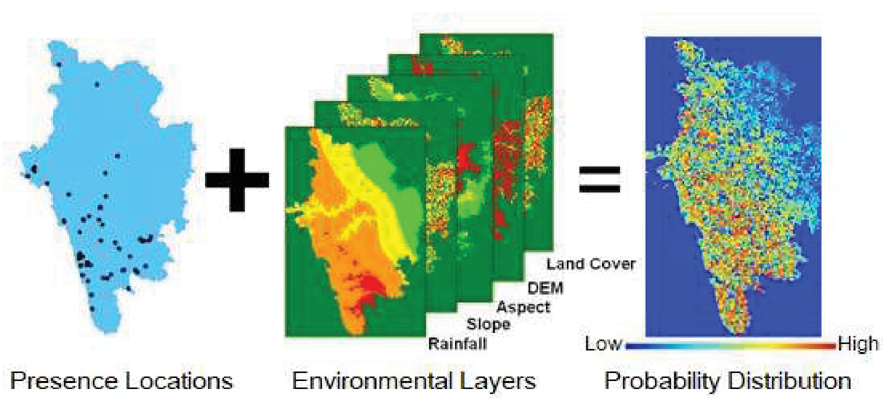

Figure A.1: Illustrated development of SDM, displaying how the species presence occurrences are added to the layered environmental variables to create a probability distribution. Adapted from *Ramachandra et al.* (2010)

A.1 Maximum Entropy

Maximum entropy refers to information entropy, which is the quantitative measure of uncertainty of the outcome of a draw from a probability distribution (*Harte and Newman*, 2014). If a probability distribution is completely uniform (ie. no areas of higher or lower probability), the likelihood of getting a particular outcome from a draw from the distribution is $1/n$, where n is equivalent to the number of possible outcomes. This relationship can be represented by the uniform distribution equation:

$$p_i = \frac{1}{n} \quad \text{for all} \quad i \in \{1, ..., n\}$$

Therefore any one outcome is equally as likely to occur as any other outcome at any point on the distribution, which means that the uncertainty in obtaining a particular outcome from a draw is at its maximum. This maximum uncertainty directly translates to maximum entropy.

MaxEnt finds the probability distribution of a species through a maximum entropy probability distribution that has been subjected to constraints, which represent the patterns in the environmental variables that correspond with species presence. MaxEnt requires that the predicted model fulfil these relationships (*Merow et al.*, 2013). For example, the constraints ensure that the mean empirical value of Variable X in presence locations is close (within set error bounds) to the mean modelled value of Variable X in the predicted presence locations (*Elith et al.*, 2010). The constraints ensure that the environmental conditions that empirically correspond with species presence are replicated throughout the model as indicators of predicted presence.

Many distributions could fit the constraints applied to the model by the environmental variables but they would all require additional information or assumptions about the species distribution. Since maximum entropy is a highly uninformed distribution, the only assumption it makes about the species distribution is the complete uncertainty in presence outcome. Thus maximum entropy is used to fit the constraints in order to avoid false assumptions. In sum, MaxEnt functions by modelling a maximum entropy probability distribution and moving away from that only as much as it is forced to in order to accommodate the constraints (*Harte and Newman*, 2014).

A.2 Output Formats

MaxEnt offers four possible output formats for the model: raw, cumulative, log, and complementary log-log (cloglog). However, I will only explain here the two formats that were used in my thesis work – raw and cloglog.

The unit used in the raw output format is Relative Occurrence Rate, or ROR, which is the relative likelihood of species presence in one raster cell as compared to another (*Phillips et al.*, 2006). It is the likelihood that a randomly chosen presence comes from the raster cell in question. All of the raw values in a single model sum to unity and are therefore scale dependent; they must sum to one no matter the size of the area being modelled, so a larger area will produce lower values than a smaller area (assuming equal cell sizes between the two areas) (*Phillips and Dudík*, 2008). The raw values are often very small in order to sum to unity and this can make their interpretation and use difficult (*Baldwin*, 2009).

The unit for the cloglog output format is Probability Of Presence, or POP, which is the absolute likelihood of species presence within a raster cell. POP is calculated by applying to the raw output a transformation involving the value c, which is the ratio of the species' true abundance to the abundance predicted by the model. The cloglog POP value which corresponds to a raw ROR value of r is:

$$1 - \exp(-cr)$$

which provides an estimate of probability of presence between 0 and 1 (*Phillips et al.*, 2017; *Phillips*, 2017). The default value of c is 0.632 and while it is not arbitrary, it is based on many assumptions, including the assumption that a typical presence has an expected abundance of one individual per sampling quadrat. The true value of c for each map made in MaxEnt is unknown and heavily depends on details of the sampling design (*Phillips*, 2017).

The raw model (ROR) compares raster cells in a relative manner, calculating how much higher or lower the likelihood of presence is between cells but without grounding those relative measures in an absolute value, while the cloglog model (POP) calculates absolute probability values for the cells through the transformation of the raw values via an assumed value, c. Hence ROR is not equivalent to POP and it is problematic to use the terms interchangeably. However, both output types have their advantages and disadvantages and the selection of output format should depend on the end-user and the future use of the model.

A.3 ROC Curves

Receiver operating characteristics (ROC) curves illustrate the predictive ability of a binary presence/absence model. They graphically display the comparison of known points with predicted results and can be used to determine the efficacy of a model. This comparison is done through the True Positive Rate (TPR, or sensitivity) and the False Positive Rate (FPR, or $1-$specificity), which calculate the percentage of true or false positives predicted by the model.

MaxEnt produces ROC curves for the models it creates. However, MaxEnt uses presence-only data, not presence and absence. Since ROC curves require absences, their calculation in MaxEnt can be difficult to conceptualise. The ROC curves in MaxEnt are obtained by plotting the TPR by the FPR at various suitability thresholds. These thresholds are the response values at which habitat is defined as suitable versus unsuitable (*Phillips et al.*, 2006). For example, a threshold value of X would mean that all cells with a response value of X or lower would be considered unsuitable. For the sake of the ROC analysis, these unsuitable areas of relatively low likelihood are considered areas of predicted absence.

This suitability dichotomy is then compared to a number of test localities, which are points within the modelled space made up of both the presence points and randomly selected background (or pseudo-absence) points. This comparison allows the creation of an error matrix that compares the predicted presence and absence (ie. the suitable and unsuitable modelled areas) to the known presence and absence (ie. the presence and background localities) (Table A.1). The matrix provides TPR and FPR values for that particular threshold, which can be plotted against each other to provide a point on a ROC curve. This process is repeated for multiple threshold values, creating a series of points that are then connected into an ROC curve (Figure A.2).

The absences involved in the making of the ROC curves are actually pseudo-absences and the interpretation of the curves assumes that limitation. The ROC's area under the curve (AUC) is a useful metric by which to compare model performance to both random chance (AUC = 0.5) and other versions of the model (*Phillips et al.*, 2006). A higher AUC implies better model performance compared to reality. The exact maximum AUC of any presence-only ROC curve is unknown, making absolute statements of performance difficult (*Phillips et al.*, 2006).

Table A.1: An example of an error matrix used in the development of ROC curves. It compares presence and background localities (Known) to suitable and unsuitable modelled areas (Predicted). The TPR is the proportion of correctly predicted presence localities, or $\frac{w}{w+y}$, and the FPR is the proportion of falsely predicted presence localities, or $\frac{x}{x+z}$.

Predicted	Known	
(Model)	Presence	Absence
Presence	w	x
Absence	y	z

A.4 Response Curves

MaxEnt produces response curves to demonstrate the modelled response of a species to the change of an environmental variable, where response is a universal term for ROR and POP (and, in the case of the cumulative output format, percentage). The response is displayed as a curve for the continuous environmental variables, and as histograms for the categorical variables.

For each environmental variable in the model, two response curves are created. The first represents the marginal response, which is a measure of the species response to the variable being considered (main variable) when the other environmental variables (background variables) are kept at their average sample value. The second curve represents the singular response, which is a model where only the main variable is included, ignoring the effect of the background variables.

Substantial differences in the responses between the first and second curves are indicative that the variables are dependent on each other. In Figure A.3, the change between the marginal response (red) and the singular response (blue) in category C implies that the background variables induce a considerable effect on the species response – their presence or absence in the response calculation produces a large difference in the response outcome. This suggests that the background variables have response effects that are somehow confounded with the main variable. The magnitude in the response differences of the marginal variable versus the singular variable is also of note. In Figure A.3, the marginal response was similar across all three categories, whereas the singular response varied substantially between the categories. This implies that the effect that the main variable has on the species response by itself is overwhelmed by the effect of the background variables, suggesting that it is not a critical variable in modelling response.

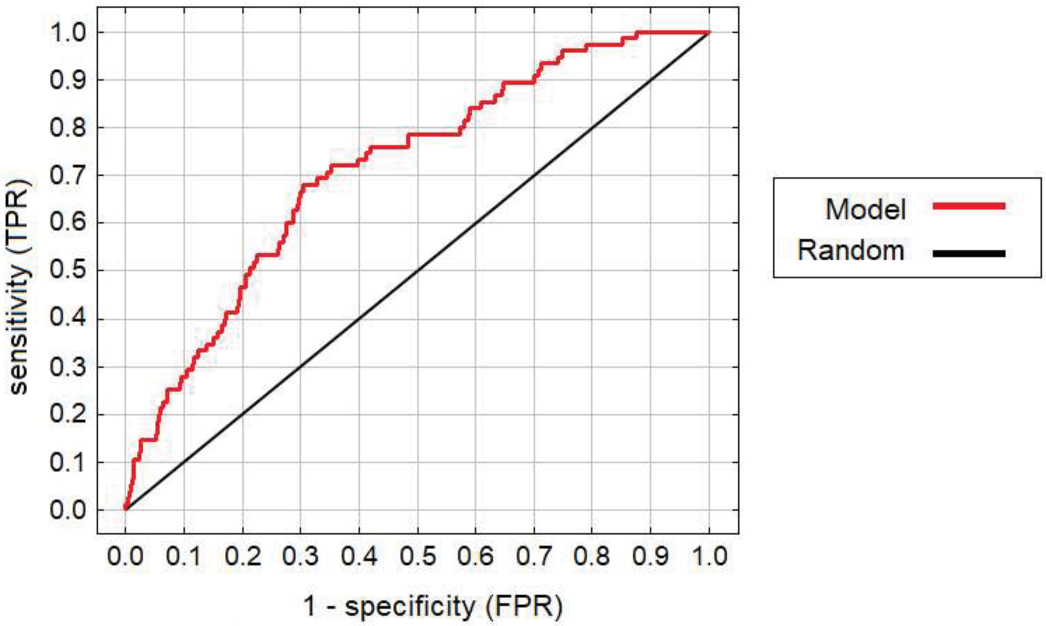

Figure A.2: Example of an ROC curve created in MaxEnt. The y-axis is the TPR and the x-axis is the FPR. The red line is a series of points, each representing the TPR/FPR at a single suitability threshold, connected together. It represents the performance of the model. The model line is higher than the black line, which represents the random model with an AUC of 0.5; therefore, the model performs better than random and has an AUC of higher than 0.5.

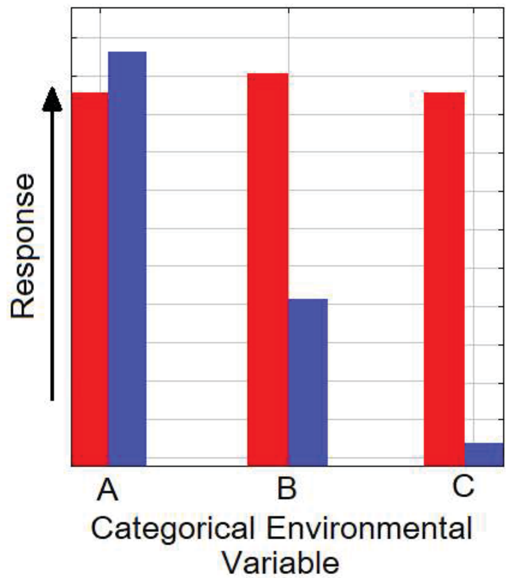

Figure A.3: Example of a species response curve for a categorical environmental variable. The red bars represent the marginal response to the main variable with the background variables included as averages. The blue bars represent the singular response of the main variable. The y-axis is the response of the species and the x-axis is the categories of the main environmental variable.

REFERENCES

Able, K. W., K. L. Heck, M. P. Fahay, and C. T. Roman, Use of salt-marsh peat reefs by small juvenile lobsters on Cape Cod, Massachusetts, *Estuaries*, *11*, 83–86, 1988.

Acheson, J., *Capturing the Commons: Devising Institutions to Manage the Maine Lobster Industry*, University Press of New England, 2003.

Aiken, D., and S. Waddy, Environmental influence on recruitment of the American lobster, *Homarus americanus*: A perspective, *Canadian Journal of Fisheries and Aquatic Sciences*, *43*, 2258–2270, 1986.

Aller, R. C., *Nitrogen Cyclings in Coastal Marine Environments*, chap. Chapter 13: Benthic Fauna and Biogeochemical Processes in Marine Sediments: The Role of Burrow Structures, pp. 301–338, John Wiley & Sons Ltd., 1988.

Anderson, J. T., R. S. Gregory, and W. T. Collins, Acoustic classification of marine habitats in coastal Newfoundland, *ICES Journal of Marine Science*, *59*, 156–167, 2002.

Anderson, J. T., D. V. Holliday, R. Kloser, D. G. reid, and Y. Simard, Acoustic seabed classification: current practice and future directions, *ICES Journal of Marine Science*, *65*, 1004–1011, 2008.

Baden, S. P., and D. M. Neil, Manganese accumulation by the antennule of the norway lobster *Nephrops norvegicus* (l.) as a biomarker of hypoxic events, *Marine Environmental Research*, *55*, 59–71, 2003.

Baldwin, R. A., Use of maximum entropy modeling in wildlife research, *Entropy*, *11*, 854–866, 2009.

Barrell, J., J. Grant, A. Hanson, and M. Mahoney, Evaluating the complementarity of acoustic and satellite remote sensing for seagrass landscape mapping, *International Journal of Remote Sensing*, *36*, 4069–4094, 2015.

Barshaw, D. E., and D. R. Bryant-Rich, A long-term study on the behavior and survival of early juvenile American lobster, *Homarus americanus*, in three naturalistic substrates: eelgrass, mud, and rocks, *Fisheries Bulletin*, *86*, 789–796, 1988.

Barshaw, D. E., K. W. Able, and J. Kenneth L. Heck, Salt marsh peat reefs as protection for postlarval lobsters *Homarus americanus* from fish and crab predators: comparisons with other substrates, *Marine Ecology Progress Series*, *106*, 203–206, 1994.

Berrill, M., and R. Stewart, Tunnel-digging in mud by newly-settled American lobsters, *Homarus americanus*, *Journal of Fisheries Research Board of Canada*, *30*, 285–287, 1973.

Bertness, M. D., Predation, physical stress, and the organization of tropical rocky intertidal hermit crab community, *Ecology*, *62*, 411–425, 1981.

BioSonics, Hydroacoustic assessment workshop, *Tech. rep.*, Biosonics Inc., 2015.

Bologna, P. A. X., Kelp beds as habitat for American lobster *Homarus Americanus*, *Marine Ecology Progress Series*, *100*, 127–134, 1993.

Botero, L., and J. Atema, Behavior and substrate selection during larval settling in the lobster *Homarus americanus*, *Journal of Crustacean Biology*, *2*, 59–69, 1982.

Bridges, C. R., and A. R. Brand, Oxygen consumption and oxygen-dependence in marine crustaceans, *Marine Ecology Progress Series*, *2*, 133–141, 1980.

Brown, C. J., A. Mitchell, D. S. Limpenny, M. R. Robertson, M. Service, and N. Golding, Mapping seabed habitats in the Firth of Lorn off the west coast of Scotland: evaluation and comparison of habitat maps produced using the acoustic ground-discrimination systems, *RoxAnn*, and sidescan sonar, *ICES Journal of Marine Science*, *62*, 790–802, 2005.

Brown, C. J., S. J. Smith, P. Lawton, and J. T. Anderson, 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, 2011.

Cobb, J. S., and G. R. Tamm, Social conditions increase intermolt period in juvenile lobsters *Homarus americanus*, *Journal of Fisheries Research Board Canada*, *31*, 1941–1943, 1974.

Cobb, J. S., and R. A. Wahle, Early life history and recruitment processes of clawed lobsters, *Crustaceana*, *67*, 1–25, 1993.

Cobb, J. S., T. Gulbransen, B. F. Phillips, D. Wang, and M. Syslo, Behavior and distribution of larval and early juvenile *Homarus americanus*, *Canadian Journal of Fisheries and Aquatic Sciences*, *40*, 2184–2188, 1983.

Coffen-Smout, S., D. Shervill, D. Sam, C. Denton, and J. Tremblay, Mapping inshore lobster landings and fishing effort on a Maritimes Region modified grid system, *resreport 3024*, Fisheries and Oceans Canada, 2013.

Cooper, R. A., and J. R. Uzmann, *The Biology and Management of Lobsters Volume 2: Ecology and Management*, chap. 3, pp. 97–139, Academic Press, 1980.

Crowder, L., and E. Norse, Essential ecological insights for marine ecosystem-based management and marine spatial planning, *Marine Policy*, *32*, 772–778, 2008.

Davies, C. E., D. Moss, and M. O. Hill, EUNIS habitat classification: Revised 2004, *resreport*, European Environment Agency, 2004.

DFO, D., 1995 volume of Atlantic coast commercial landings, by region (metric tonnes, live weight), 1995.

DFO, D., 2005 volume of Atlantic coast commercial landings, by region (metric tonnes, live weight), 2005.

DFO, D., 2015 Atlantic coast commerical landings, by region (metric tonnes, live weight), 2015.

DFO, D., 2016 value of Atlantic & Pacific coasts commercial landings by province (thousand dollars), 2016.

Douvere, F., The importance of marine spatial planning in advancing ecosystem-based sea use management, *Marine Policy*, *32*, 762–771, 2008.

Elith, J., S. J. Phillips, T. Hastie, M. Dudk, Y. E. Chee, and C. J. Yates, A statistical explanation of MaxEnt for ecologists, *Diversity and Distributions*, *17*, 43–57, 2010.

Eriksson, S. P., and S. P. Baden, Behaviour and tolerance to hypoxia in juvenile norway lobster (*Nephrops norvegicus*) of different ages, *Marine Biology*, *128*, 49–54, 1997.

Fader, G. B., L. H. King, and B. MacLean, Surficial geology of the eastern Gulf of Maine and Bay of Fundy, *Geological Survey of Canada Papers*, *76*, 1977.

Foley, M. M., B. S. Halpern, F. Micheli, M. H. Armsby, M. R. Caldwell, C. M. Crain, E. Prahler, N. Rohr, D. Sivas, M. W. Beck, M. H. Carr, L. B. Crowder, J. E. Duffy, S. D. Hacker, K. L. McLeod, S. R. Palumbi, C. H. Peterson, H. M. Regan, M. H. Ruckelshaus, P. A. Sandifer, and R. S. Steneck, Guiding ecological principles for marine spatial planning, *Marine Policy*, *34*, 955–966, 2010.

Foster-Smith, R. L., C. J. Brown, W. J. Meadows, W. H. White, and D. S. Limpenny, Mapping seabed biotopes at two spatial scales in the eastern English Channel. Part 2. comparison of two acoustic ground discrimination systems, *Journal of the Marine Biological Association of the United Kingdom*, *84*, 489–500, 2004.

Franklin, J., *Mapping species distributions: Spatial inference and prediction.*, Cambridge University Press, 2010.

Freitas, R., L. Sampaio, J. Oliveira, A. M. Rodrigues, and V. Quintino, Validation of soft bottom benthic habitats identified by single-beam acoustics, *Marine Pollutant Bulletin*, *53*, 72–79, 2006.

Freitas, R., A. M. Rodrigues, E. Morris, jose Lucas Perez-Llorens, and V. Quinto, Single-beam acoustic ground discrimination of shallow water habitats: 50 kHz or 200 kHz frequency survey?, *Estuarine, Coastal and Shelf Science*, *78*, 613–622, 2008.

Freitas, R., F. Ricardo, F. Pereira, L. Sampaio, S. Carvalho, M. Gaspar, V. Quintino, and A. M. Rodrigues, Benthic habitat mapping: Concerns using a combined approach (acoustic, sediment and biological data), *Estuarine, Coastal and Shelf Science*, *92*, 598–606, 2011.

Green, E. P., P. J. Mumby, A. j. Edwards, and C. D. Clark, A review of remote sensing for the assessment and management of tropical coastal resources, *Coastal Management*, *24*, 1–40, 1996.

Hagerman, L., and R. F. Uglow, Effects of hypoxia on the respiratory and circulatory regulation of *Nephrops norvegicus*, *Marine Biology*, *87*, 273–278, 1985.

Harris, B. P., and K. D. E. Stokesbury, Shell growth of sea scallops (*Placopecten magellanicus*) in the southern and northern Great South Channel, USA, *ICES Journal of Marine Science*, *63*, 811–821, 2006.

Harte, J., and E. A. Newman, Maximum information entropy: a foundation for ecological theory, *Trends in Ecology and Evolution*, *29*, 384–389, 2014.

Holmes, K., K. V. Neil, B. Radford, G. Kendrick, and S. Grove, Modelling distribution of marine benthos from hydroacoustics and underwater video, *Continental Shelf Research*, *28*, 1800–1810, 2008.

Holthuis, L. B., *Marine Lobsters of the World: an annotated and illustrated catalogue of species of interest to fisheries known to date*, vol. 13, Food and Agriculture Organization of the United Nations, 1991.

Howard, A. E., and D. B. Bennett, The substrate preference and burrowing behaviour of juvenile lobsters (*Homarus gammarus* (L.)), *Journal of Natural History*, *13*, 433–438, 1979.

Howard, R. J., and J. S. Larson, A stream habitat classification system for beaver, *The Journal of Wildlife Management*, *49*, 19–25, 1985.

Hudon, C., Ecology and growth of postlarval and juvenile lobster, *Homarus americanus*, off Îles de la Madeleine (Quebec), *Canadian Journal of Fisheries and Aquatic Sciences*, *44*, 1855–1869, 1987.

Hutin, E., Y. Simard, and P. Archambault, Acoustic detection of a scallop bed from a single-beam echosounder in the St. Lawrence, *ICS Journal of Marine Science*, *966-983*, 2005.

ICES (Ed.), *ICES Report of the Working Group on Marine Habitat Mapping (WGMHM)*, 2017.

Ivany, R., I. d'Entremont, D. Christmas, S. Fuller, and J. Bragg, The report of the Nova Scotia commission on building our new economy, *Tech. rep.*, One Nova Scotia, 2014.

Jaynes, E. T., Information theory and statistical mechanics, *Physical Review*, *106*, 620, 1957.

Jones, H. G., and X. R. R. Sirault, Scaling of thermal images at different spatial resolution: the mixed pixel problem, *Agronomy*, *4*, 380–396, 2014.

Karnofsky, E. B., and H. J. Price, Behavioural response of the lobster *Homarus americanus* to traps, *Canadian Journal of Fisheries and Aquatic Science*, *46*, 1625–1632, 1989.

Laurel, B. J., R. S. Gregory, J. A. Brown, J. K. Hancock, and D. C. Schneider, Behavioural consequences of density-dependent habitat use in juvenile cod *Gadus morhua* and *G. ogac*: the role of movement and aggregation, *Marine Ecology Progress Series*, *272*, 257–270, 2004.

Lawton, P., and K. L. Lavalli, *Biology of the Lobster* Homarus americanus, chap. 4. Postlarval, Juvenile, Adolescent, and Adult Ecology, pp. 47–88, Academic Press, 1995.

Lecours, V., R. Devillers, D. C. Schneider, V. L. Lucieer, C. J. Brown, and E. N. Edinger, Spatial scale and geographic context in benthic habitat mapping: review and future directions, *Marine Ecology Progress Series*, *535*, 259–284, 2015.

Marshall, J., Landlords, leaseholders & sweat equity: changing property regimes in aquaculture, *Marine Policy*, *25*, 335–352, 2001.

McMahon, B. R., and J. L. Wilkens, Simultaneous apnoea and bradycardia in the lobster *Homarus americanus*, *Canadian Journal of Zoology*, *50*, 165–170, 1971.

McMahon, B. R., and J. L. Wilkens, Respiratory and circulatory response to hypoxia in the lobster *Homarus americanus*, *Journal of Experimental Biology*, *62*, 637–655, 1975.

Merow, C., M. J. Smith, and J. John A. Silander, A practical guide to MaxEnt for modeling species' distributions: what it does, and why inputs and settings matter, *Ecography*, *36*, 1058–1069, 2013.

Murphy, H. M., and G. P. Jenkins, Observational methods used in marine spatial monitoring of fishes and associated habitats: a review, *Marine and Freshwater Research*, *61*, 236–252, 2010.

Nelson, K., D. Hedhecock, W. Borgeson, E. Johnson, R. Daggett, and D. Aronstein, Density-dependent growth inhibition in lobsters, *Homarus* (Decapoda, Nephrodidae), *The Biological Bulletin*, *159*, 162–176, 1980.

Paille, N., B. Sainte-Marie, and J.-C. Brêthes, Behavior, growth and survival of Stage V lobsters (*Homarus americanus*) in relation to shelter availability and lobster density, *Marine and Freshwater Behaviour and Physiology*, *35*, 203–219, 2002.

Palma, A. T., R. A. Wahle, and R. S. Steneck, Different early post-settlement strategies between American lobsters *Homarus americanus* and rock crabs *Cancer irroratus* in the Gulf of Maine, *Marine Ecology Progress Series*, *162*, 215–225, 1998.

Petitgas, P., J. Mass, P. Beillois, E. Lebarbier, and A. L. Cann, Sampling variation of species identification in fisheries acoustic surveys based on automated procedures associating acoustic images and trawl hauls, *ICES Journal of Marine Science*, *60*, 437–445, 2003.

Phillips, S., and M. Dudík, Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation., *Ecography*, *31*, 161–175, 2008.

Phillips, S., R. Anderson, and R. Schapire, Maximum entropy modeling of species geographic distributions., *Ecological Modelling*, *190*, 231–259, 2006.

Phillips, S. J., A brief tutorial on Maxent, 2017, available from: URL: https://biodiversityinformatics.amnh.org/open_source/maxent/Maxent_tutorial2017.pdf.

Phillips, S. J., R. P. Anderson, M. Dudík, R. E. Schapire, and M. E. Blair, Opening the black box: an open-source release of Maxent, *Ecography*, *40*, 887–893, 2017.

Pineda, J., and H. Caswell, Dependence of settlement rate of suitable substrate area, *Marine biology*, *129*, 541–548, 1997.

Pottle, R. A., and R. W. Elner, Substrate preference behavior of juvenile American lobsters, *Homarus americanus*, in gravel and silt-clay sediments, *Canadian Journal of Fisheries and Aquatic Science*, *39*, 928–932, 1982.

Pulliam, H. R., On the relationship between niche and distribution, *Ecology Letters*, *3*, 349–361, 2000.

Ramachandra, T. V., U. Kumar, B. H. Aithal, P. G. Diwakar, and N. V. Joshi, Landslide susceptible locations in western Ghats: prediction through OpenModeller, in *Proceedings of the 26th Annual In-House Symposium on Space Science and Technology*, pp. 65–74, Indian Institute of Science, 2010.

Rooper, C. N., M. Zimmermann, M. M. Prescott, and A. J. Hermann, Predictive models of coral and sponge distribution, abundance and diversity in bottom trawl surveys of the aleutian islands, alaska, *Marine Ecology Progress Series*, *503*, 157–176, 2014.

Schweizer, D., R. A. Armstrong, and J. Posada, Remote sensing characterization of benthic habitats and submerged vegetation biomass in Los Roques Archipelago National Park, Venezuela, *International Journal of Remote Sensing*, *26*, 2657–2667, 2005.

Selgrath, J. C., K. A. Hovel, and R. A. Wahle, Effects of habitat edges on American lobster abundance and survival, *Journal of Experimental Marine Biology and Ecology*, *353*, 253–264, 2007.

Shaw, J., B. J. Todd, and M. Li, Seascapes, Bay of Fundy, offshore Nova Scotia/New Brunswick, *Tech. rep.*, Geological Survey of Canada, 2012, open File 7028.

Shucksmith, R. J., and C. Kelly, Data collection and mapping - principles, processes and application in marine spatial planning, *Marine Policy*, *50*, 27–33, 2014.

Sly, P. G., and W. Busch, *The development of an aquatic habitat classification system for lakes*, chap. Introduction to the process, procedure, and concepts used in the development of an aquatic habitat classification system for lakes, pp. 3–21, CRC Press, 2018.

Spanier, E., What are the characteristics of a good artificial reef for lobsters?, *Crustaceana*, *67*, 173–186, 1993.

Steneck, R., Possible demoraphic consequences of intraspecific shelter competition among American lobsters, *Journal of Crustacean Biology*, *26*, 628–638, 2006.

St.Martin, K., and M. Hall-Arber, The missing layer: Geo-technologies, communities, and implications for marine spatial planning, *Marine Policy*, *32*, 779–786, 2008.

Tang, F., T. Minch, K. Dinning, C. J. Martyniuk, R. Kilada, and R. Rochette, Size-at-age and body condition of juvenile American lobsters (*Homarus americanus*) living on cobble and mud in a mixed-bottom embayment in the Bay of Fundy, *Marine Biology*, *162*, 69–79, 2015.

Tremblay, M. J., S. J. Smith, B. J. Todd, P. M. Clement, and D. L. McKeown, Associations of lobsters (*Homarus americanus*) off southwestern Nova Sscotia with bottom type from images and geophysical maps, *ICES Journal of Marine Sciences*, *66*, 2060–2067, 2009.

Turner, C., Using habitat suitability modelling to assess the relationships between shrimp trawling and the distributions of deep sea corals, mathesis, Imperial College London, 2014.

VanOlst, J. C., J. M. Carlberg, and R. F. Ford, Effects of substrate type and other factors on the growth, survival, and cannibalism of juvenile *Homarus americanus* in mass rearing systems, *Journal of the World Aquaculture Society*, *6*, 261–274, 1975.

Wahle, R., and R. Steneck, Recruitment habitats and nursery grounds of the American lobster *Homarus americanus*: A demographic bottleneck?, *Marine Ecology Progress Series*, *69*, 231–243, 1991.

Wahle, R. A., Substratum constraints on body size and the behavioral scope of shelter use in the American lobster, *Journal of Experimental Biology and Ecology*, *159*, 59–75, 1992.

White, W. H., A. R. Harborne, I. S. Sotheran, R. Walton, and R. L. Foster-Smith, Using an acoustic ground discrimination system to map coral reef benthic classes, *International Journal of Remote Sensing*, *24*, 2641–2660, 2003.

Whitlatch, R. B., Seasonal changes in the community structure of the macrobenthos inhabiting the intertidal sand and mud flats of Barnstable Harbor, Massachusetts, *Biology Bulletin*, *152*, 275–294, 1977.

Wiber, M. G., S. Young, and L. Wilson, Impact of aquaculture on commercial fisheries: Fisherman's local ecological knowledge, *Human Ecology*, *40*, 29–40, 2012.