Distribution of Queen conch (*Strombus gigas*) on the Pedro Bank, Jamaica: Descriptive and Predictive distribution models

Ricardo Anthony Morris

Faculty of Engineering and Natural Sciences
University of Iceland
2016

Distribution of Queen Conch (*Strombus gigas*) on the Pedro Bank, Jamaica: Descriptive and predictive distribution models

Ricardo Anthony Morris

90 ECTS thesis submitted in partial fulfillment of a
Magister Scientiarum degree in Marine Biology

Advisor(s)
Jörundur Svavarsson
Julian Burgos

Faculty Representative
Halldór P. Halldórsson

Faculty of Life and Environmental Sciences
School of Engineering and Natural Sciences
University of Iceland
Reykjavik, March 2016

Distribution of Queen conch (*Strombus gigas*) on the Pedro Bank, Jamaica: Descriptive and Predictive distribution models
90 ECTS thesis submitted in partial fulfillment of a *Magister Scientiarum* degree in Biology

Abstract

Species distribution models (SDM) are useful tools for describing and predicting species ecological role within its community or ecosystem. They are increasingly becoming important in the context of marine resource management and conservation in light of the the relative difficulty and expense of obtaining quality marine biological and environmental data. SDM have been applied to the marine realm in areas such as marine spatial planning (MSP), prioritizing for the establishment of protected areas, predicting and planning for the impact of non-native species and climate change mitigation. There is an on-going effort by private and public stakeholders to further develop and implement ecosystem based management approaches to Jamaica's marine resources, in particular on socio-economically important species and habitats. In this context descriptive and predictive distribution models were developed for different size and age groups of the commercially important gastropod mollusc, Queen conch (*Strombus gigas*), on the Pedro Bank Jamaica. Species occurrence data from four abundance surveys (2002, 2007, 2011 and 2015) were modelled against the environmental variables; depth, substrate and primary production using generalized additive models (GAM) and Maxent. Descriptive results revealed that for both juveniles and adults there is a higher tendency toward shallower depths and substrates with relatively high amounts of their marcoalgal food. However, it is substrate complexity rather than specific substrate types that may be most important. Predictive models showed that not all supposedly suitable areas were being occupied by the species, but importantly also suggest priority areas for management of the species and its habitat in the context of the Pedro Banks' increasingly multi-use nature.

Key words: *Strombus gigas*, realized niche, potential niche, Pedro Bank, species distribution models, biological variables, environmental variables, marine spatial planning

Útdráttur

Útbreiðslulíkön tegunda (Species distribution models, SDM) eru hentug tæki til að lýsa og spá fyrir um útbreiðslu tegunda og hlutdeild þeirra og stöðu í samfélögum og vistkerfum. Notkun þeirra hefur aukist verulega við stjórnun sjávarauðlinda og verndun þeirra, einkum í ljósi erfiðleika og mikils kostnaðar við að afla magnbundinna gagna, bæði líffræðilegra og umhverfisgagna. Slíkum líkönum hefur verið beitt á lífverur og umhverfi sjávar, m.a. við skipulagningu sjávarsvæða, skipulagningu verndaðra svæða, til að meta áhrif aðkomutegunda og við mat á áhrifum loftslagsbreytinga. Á Jamaíku er nú aukin áhersla á að beita vistfræðilegum aðferðum við auðlindastjórnun, einkum vegna nytjategunda og mikilvægra búsvæða í sjó. Í tengslum við þetta voru lýsandi líkön (descriptive) og spálíkön (predictive) útbúin fyrir mismunandi strærðardreifingu og aldursflokka snigilsins tröllajöfurs (Queen conch, *Strombus gigas*) á Pedro Banka á Jamaíku, en tegundin er mikilvæg nytjategund þar. Gögn um magn og útbreiðslu tegundarinnar úr fjórum sýnatökum (2002, 2007, 2011 og 2015), ásamt umhverfisbreytunum dýpi, gerð undirlags og frumframleiðslu, voru sett í GAM (Generalized additive models) og Maxent líkön. Niðurstöður úr lýsandi líkönum sýndu tilhneigingu til þess að bæði ungviði og fullorðin dýr héldu sig frekar á grynnra vatni og á undirlagi með hlutfallslega miklu magni af stórþörungum, sem eru helsta fæða tröllajöfurs. Hins vegar bendir ýmislegt til meira mikilvægis fjölbreytileika undirlagsins, frekar en sérstakrar undirlagsgerðar. Spálíkönin sýndu að tegundin finnst ekki víða þar sem ætla má að kjöraðstæður séu fyrir hana. Út frá líkönunum má skilgreina forgangssvæði vegna auðlindastjórnunar á tröllajöfri, en slíkt er mikilvægt í ljósi aukinnar og fjölbreyttari nýtingar lífríkisins á Pedro Banka.

Table of Contents

List of Figures

List of Tables

Acknowledgements

1 Introduction

1.1 Conch benthic distribution and life history

Queen conch (*Strombus gigas* Linnaeus, 1758; synonym *Lobatus gigas* (Linnaeus, 1758)) is a large gastropod mollusc (Figure 1) native to shallow waters usually less than 50 m of the western Atlantic region, ranging from Bermuda in the north down to the Gulf of Mexico, the greater Caribbean region, and the coast of northern South America (Chakallal and Cochrane, 1996). The animal often grows to around 25 cm in shell length which it attains after 3 to 4 years and at which time most individuals are also sexually mature (Brownell and Stevely, 1981; Prada *et al*., 2008).

Figure 1. Queen conch viewed from different angles; above view (left), underside view (centre) and lateral view (right).

The species is an important part of the comminities and ecosystems in which it occur often being one of the primary benthic grazing herbivores, feeding on marcoalgae and detritus (Brownell and Stevely, 1981). The species is also a food source for a number of other species ranging from planktonic crustaceans and fish which feed on the young planktonic life stages to larger marine animals such as Nurse Sharks (*Ginglymostoma cirratum* Bonnaterre, 1788), Spiny lobsters (*Palinurus argus* Latreille, 1804) and octopuses (*Octupus vulgaris* Lamarck, 1798) which prey on benthic life stages (Prada *et al*., 2008).

Like most marine species, their distribution over space and time is determined by a complex relationship among various biological and environmental variables including those of anthropogenic origin (Brownell and Stevely, 1981; Prada *et al*., 2008). The extent, composition and nature of the substratum is one of the key determinants of the distribution of benthic marine organisms (Reiss *et al*., 2014), and such is the case with the Queen conch. Throughout its benthic life history and ontogeny Queen conch inhabit mainly flat, shallow sand-based or hard bottom (weathered dead coral reef) habitat associated with relatively high primary production needed for their main macroalgae and seagrass detritus food sources (Brown and Stevely, 1981; Stoner *et al.*, 1996b). Their distribution, like that of many other benthic species, is therefore closely related to access to and availability of

their food resulting in the species occurring often in clusters throughout their range (Stoner and Lally, 1994).

Different substrate types or combination of substrate types including coral reefs, reef pavement, coral rubble (weathered coral reef fragments and biogenic debris), sands of various grain sizes, and seagrasses are thought to influence the abundance and population structure of these clusters (Brownell and Stevely, 1981). This is partly due to their differing affinity for growth of their marcoalgal food and for providing suitable spawning and nursery habitat (Brownell and Stevely, 1981; Stoner and Sandt, 1992). As a result different age and size groups display daily and seasonal movements between different feeding grounds as well as movement between feeding grounds and spawning grounds in the case of sexually mature adults (Stoner and Sandt, 1992). Daily movements for food has been estimated to range between 2 and 15 m/day while an individuals' annual range, including for reproductive activity, is thought to range between 2 ha/year (0.02 km²) and 30 ha/year (0.3 km²) (Stoner and Ray-Culp, 2000).

There is significant movement of the species among different substrates for the purpose of reproduction (Stoner and Sandt, 1992). Sexually mature conch of around 3 years and older will move from marcoalgal hard substrates (coral reefs, coral rubble and reef pavement) where they feed during the months of May to June (Stoner and Sandt, 1992). They then move to flat sandy habitat where they copulate and spawn during the warmer months of the year peaking during the period of July to August (Stoner and Culp, 2000). Very little feeding occurs during the reproductive season (Stoner and Sandt, 1992). At the end of the reproductive season, roughly from August to September, they migrate to hard substrate where they feed until the next reproductive season (Stoner and Sandt, 1992).

The migration of juveniles and sexually mature adults are thought to produce a type of ontogenic stratification in some parts of the Caribbean region particularly based on substrate type and depth (Stoner and Ray-Culp, 2000). In these areas, smaller juvenile conch tend to occur more often in shallower waters associated with abundant macroalgae and seagrass where they aggregate in high densities as a mechanism to avoid predation and increase overall survivorship (Stoner and Lally, 1994). Larger adults, on the other hand, are more numerous in deeper waters associated with flat, less complex habitat where they may be more exposed to predation (Stoner and Sandt, 1992). This type of stratification is however not the rule as the size and composition of Queen conch aggregations are dependent on a complex of substrate types in an area rather than any one or two specific types thus both juveniles and adults may occupy the same general areas (Stoner and Lally, 1994). Juveniles may occur in deeper waters possibly as a means of avoiding mechanical disturbance and dealing with limited amounts of suitable habitat and resources (Marshak *et al*., 2006), while adults may utilize the often food-rich shallower areas as well (Stoner and Sandt, 1992)

It is therefore likely that the species' distribution is dependent on a more complex set of ecological and environmental processes which may be difficult to measure but are influential shaping distributional patterns (Stoner, 2003). These include for instance different species interactions in determining the spatial extent of nurseries, feeding grounds and spawning areas (Marshak *et al*., 2006). Other important factors include ocean currents which have important role in larval retention and settlement as the species planktonic life phase may last for up to three weeks which is ample time for wide distribution by currents

(Stoner *et al*., 1996a). Additionally, photoperiod or the relative amount of daylight per day is also thought to provide physiological stimuli for commencement of reproductive migration in sexually mature adults and also for pelagic larvae to move to the benthos to settlement (Stoner and Sandt, 1992; Stoner, 2003). This research paper will however not consider factors such as these but focus mainly on modeling direct and indirect environmental factors affecting the distribution of the benthic life stages of the species namely; depth, substrate and chlorophyll-a concentration which is used here as an indirect measure of primary productivity.

1.2 Conch fishery on the Pedro Bank

An important consideration when looking at Queen conch on the Pedro Bank is the fact that it is an important commercial species. Exports originating from the bank during the last decade and a half has averaged around 500 tonnes (MT) annually (Fisheries Division, 2013), and contributing to a regional fishery worth in excess of an estimated 60 million United States dollars annually thus underlining its tremendous socioeconomic importance to coastal communities in Jamaica and other countries in the region (Chakallal and Cochrane, 1996). That said however, by its very nature the commercial exploitation of Queen conch is a source of potential threat to the species itself and to the wider biodiversity on the bank necessitating prudent management.

The Queen conch fishery grew exponentially from the mid 1980's to the late 1990's where in fact before 1988 Jamaica averaged around 50 MT in annual exports and by 1995 exports this had risen to over 2,000 MT (Aiken *et al*., 1999). This excessive and unsustainable removal of individuals was thought to have had significant negative impacts on the distribution and population structure of the stock on the Pedro Bank, as well as causing ecological imbalances (Tewfik and Appeldoorn, 1998). Luckily through the introduction of a comprehensive management regime in the early 1990's the fishery underwent a period of contraction which brought both level of exploitation and the number of fishers down to more sustainable levels (Smikle and Appeldoorn, 2003).

The management regime implemented during this time contained a number of tools and instruments which allowed for a recovery on the path of sustainability. Perhaps the most important of these management instruments came about with the listing of the species under appendix II of the Convention on International Trade in Endangered Species (CITES) in 1992 in response to the high levels of over-exploitation and poor management which was occurring simulataneously throughout the region (Aiken, *et al*., 1999). Appendix II of the CITES lists species not yet endangered but are at risk to becoming endangered due to their commercial exploitation and obligates countries invloved in their trade to take action and prove that their level of exploitation is not detrimental to the species (Aiken, *et al*., 2006). Jamaica's response to this and the general need to improve the management of the species included the drafting of a comprehensive fishery management plan outlining the policy direction, specific management actions and the roles of various regulators involved in the fishery (Fisheries Division, 2014).

A number of initiatives within the framework of the draft management plan were implemented inluding; a fisher reduction programme, a limited entry policy and a non-transferable individual quota and total allowable catch (TAC) system which were

determined by a series of regular population surveys and other ecosystem-based research on the species and its associated habitat (Aiken, *et al*., 2006). Additionally there were a number of important pieces of legislations passed aimed at empowering regulators including regulations and amendments to the existing Fishing Industry Act of 1975 as well as new ones such as the the Aquaculture, Inland, Marine Products and By-Products Act of 1999. As a result of these and other initiatives by 2011 the Jamaican fisheries authorities were reporting that the key indicators of stock health, densities and biomass, were at levels similar to the expansion phase in the early 1990's (Fisheries Division, 2013; CRFM, 2012).

Notwithstanding the success of this management intervention it is important to understand that commercial fishing may have still had a significant impact on the species' distribution (Stoner and Ray, 1996). Stoner and Sandt (1992) as well as Stoner and Schwarte (1994) from their studies in the Bahamas discussed significant differences observed in the distribution and densities composition of conch communities in fished versus non-fished areas on substrates at different depths. Also there is the threat of localized extinction and negative population growth which can result from fishing as fishers will more likely target high density areas. This practice increase the likelihood of allee effects where reproductive adult densities become too low (~56 adult conch/ha) for effective copulation and reproduction to occur resulting in reduced stocks (Stoner and Ray-Culp, 2000). Of concern as well is the effect of the practice of removing meat from the shell underwater which results in numerous dead conch shells lining the seafloor (Tewfik and Appeldoorn, 1998; Kaiser *et al*., 2003). Live conch are thought to avoid areas where there are dead conch, potentially reducing the amount of suitable habitat for the species (Aiken *et al*., 2006).

1.3 Species distribution models

Species distribution models (SDM) are a group of tools and techniques used to describe and predict spatial patterns of species based on specific environmental parameters (Elith and Leathwick, 2009). SDMs in this sense refer to statistical and machine learning ecological models rather than theoretical and heuristic models (Guisan *et al*., 2002). These have become an important part of ecosystem-based management (ESBM) approaches all over the world, mostly in the terrestrial realm but are becoming increasingly used in the marine environment as well (Reiss *et al*., 2014). This increased usage and importance has occured especially over the last two decades as a result of a number of technical advancements in the mathematics applied to ecological data as well as increased access to more powerful computers and software able to carry out such analyses (Guisan *et al*., 2002). Ecologistist and other natural resource managers therefore, more than ever, have a greater ability to quantify, visualize and assess species-environment relationships which can then be fed into larger decision-making processes, thus leading to improved capacity to sustainably manage sensative or threatened ecosystems (Robinson *et al.,* 2011).

Management and conservation of marine ecosystems and their resources are among the most promising areas of application for SDMs especially given the often prohibitive costs associated with obtaining detailed and quality biological and environmental information (Reiss *et al*., 2014). This high cost is perhaps the greatest obstacle to SDMs' greater use in a marine context, and thus renders it still very much a developing field as data is often absent, sparse or of low quality for use in SDMs (Guisan and Zimmermann, 2000). Use of

SDMs in marine environments as a result often follow closely the pattern of easily-to-accessible or low-cost data usually in and around shallow coastal marine areas or areas having high value species or habitats (Elith and Leathwick, 2009; Reiss *et al*., 2014).

Many SDMs, particularly predictive ones, can be useful in filling knowledge gaps in relatively data-poor marine environments as they are able to model species distributions over large geographical areas with relatively small amounts of biological and environmental data input (Guisan *et al*., 2002). This attribute has lead to successful application in areas such as; (i) marine spatial planning (MSP), (ii) designing of marine species and resource monitoring programmes, (iii) management of non-indeginous species, and (iv) providing prediction of species distributions in light of various threats such as habitat loss and climate change (Robinson *et al*., 2011; Reiss *et al*., 2014). These applications have lead in many cases to better resourse management and, importantly, better management of conflicts among a growing number marine stakeholders around the world (Baldwin *et al*., 2014; Reiss *et al*., 2014).

Two of the more commonly used groups of SDMs are; (i) generalized regression methods and (ii) machine learning methods (Elith and Leathwick, 2009). Regression methods include generalized linear models (GLM) and generalized additive models (GAM) which have proven very useful in modelling species distributions of both terrestrial and marine species due to their relative simplicity, ease of interpretation and relatively low data requirement (Guisan and Zimmermann, 2000; Huang *et al*., 2011). Many machine learning SDMs have been relatively recent in their developed and include methods such as Maximum Entropy (Maxent) (Phillips *et al*., 2006), BIOCLIMatic (BIOCLIM) (Nix, 1986), and DOMAIN (Carpenter *et al*., 1993) among others. Both generalized regression methods and machine learning methods may be used to simulate species distributions that estimate different aspects of a species' specific ecological role in its community or ecosystem; that is, its realized niche and fundamental/potential niches. A species' fundamental niche refers to its distribution (or other response) as a function of physiological and ecosystem constraints, and is estimated based on theoretical concepts (Guisan and Zimmermann, 2000). The realized niche on the the other hand incorporates further contraints such as biotic interactions and competitive exclusion, and is often based on field observations (Malanson *et al*., 1992).

The success of GLMs and GAMs in distribution is largely due to the fact that they are flexible enough to allow for a multitude of distributions, correlations, variance and error structures that are more suited to ecological data (Hastie and Tibshirani, 1990; Guisan *et al*., 2002). They therefore are able to provide a truer picture of the species-environment relationship than tranditional linear regression models (LR) for instance that are bounded by rigid assumptions about the data (Zuur *et al*., 2010). These assumptions often lead to instances where ecological data are forced into a model through data transformation and other means that can lead to the loss of important data characteristics (Zuur *et al*., 2009). Perhaps the most important advantage of GLMs and GAMs is that they are able to model species abundances as opposed to presence/absence models that ignore relative abundance (Reiss *et al*., 2014).

Maxent is a presence/absence model but is recognized as having one of the best predictive performance among machine learning methods and is also well suited to the data-limited marine environments as it is robust against small sample sizes (Elith and Leathwick, 2009;

Reiss *et al.,* 2014). Maxent and other presence/absence models are often used in indicating suitability of habitats given the background and assumption of a species being present in in certain specified environmental conditions (Phillips *et al*. 2006). It is cost-effective as it essentially only require occurence (presence) records as input along with spatial environmental layers, many of which have become available as open-source datasets (Elith and Leathwick, 2009; Magris and Déstro, 2010). The predictive output, even with small amounts of input data, can give useful insight in a relatively easily understood manner into salient biological characteristics and environmental conditions affecting species distributions which otherwise would have only been speculative (Robinson *et al*., 2011). It should also be noted that the quality of output given by Maxent and any other SDM is dependent to a large extent on deligent consideration by the researcher in terms of; an appropriate sample, choosing of only relevent environmental predictors to explain a species' response, and adherence to the modelling processes including model evaluation and interpretation of its results (Reiss *et al*., 2014).

There are also a number of more general numerical approaches may also be useful in the distribution modelling process (Reiss *et al*., 2014). Heuristic clustering methods for instance can allow for the detection of discountinuous groupings (clusters) within species group, among groups of species as well as environmental variables which can be critically important information for ecologists looking to model the distribution of species (Bocard *et al*., 2011). In addition, ordination methods such as principal component analysis (PCA) and canonical correspondence analysis (CCA) are also useful tools for identifying important patterns in the relationship among modelled species and environmental data (Moore *et al*., 2010).

The focus of this research paper will be to use GAMs and Maxent to model the distribution of Queen conch in waters of the Pedro Bank, Jamaica. GAMs will be used to model the species' abundance and distribution assuming that the variables included in the models are most important in determining its' abundance and spatial extent. The Maxent method will be used to model habitat suitability for Queen conch on the bank given the specifc environmental factors.

The use of Maxent here is due to its wide-scale usage and popularity particularly in data-poor marine environments. It was thought, therefore, that for a small island developing state such as Jamaica it would be worthwhile to constrast its output against that of the GAM. This would allow for its exploration it as a frequent, viable and cost-effective addition (or alternative if needs be) to the costly abundance survey used for studying the Queen conch distribution. Implicit in these modelling exercises are the further objectives of determining the relative importance of each environmental variable in addition to providing management recomendations based on the models' outcomes.

2 Background

2.1 Study site description

The Pedro Bank is an offshore proximal bank located within the territorial waters of Jamaica roughly 80 km south of the main island (Figure 1) (Allen and Webber, 2013). The country itself is located in the western Caribbean region and is an archipelagic state consisting of the main island as well as a number of nearshore and offshore cays, shoals and banks of which the Pedro Bank is the largest and most important (Munro and Thompson, 1983). The bank is amongst the country's most valuable marine resources generating income and serving the purposes of many stakeholders and with still greater potential for generating income (Baldwin *et al*., 2014) It is essentially an underwater plateau with its highest points to the southeast, then gradually slope toward the north to around 70 m before finally descending into deep waters (Smikle and Apeldoorn, 2003). The bank also extends some 184 km west to east and about 84 km north to south covering an area of approximately 8,040 km² (804,000 ha) (Munro and Thompson, 1983).
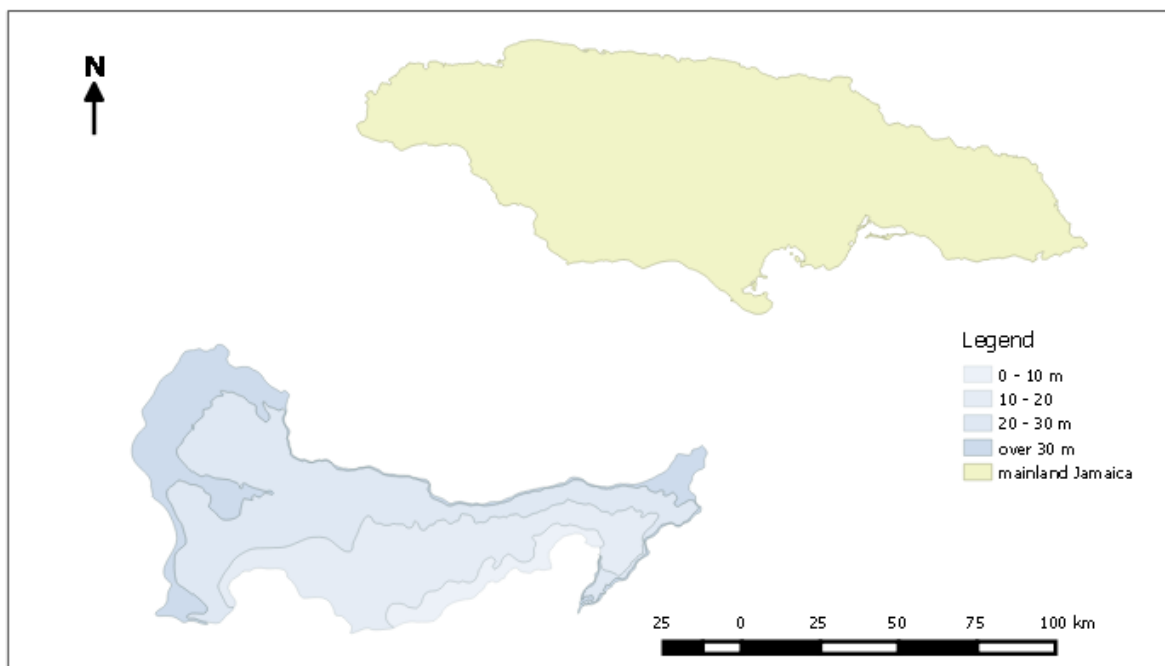
Figure 2. Map of mainland Jamaica and the Pedro Bank. Map adapted from Fisheries Division (2013) and produced using QGIS (QGIS Development Team (2015).

On its southeastern edge, the bank has a chain of four islands, the Pedro Cays, which are often used as a base for various users including; wildlife conservation interests, research and education, transport, marine resource exploration, the military, tourism, and fisheries (Allen and Webber, 2013). These in part have fostered the development of the bank over a

number of years into an important multi-use marine space serving the interest of many sectors and playing an ever increasing role in the country's socio-economic development (Baldwin *et al*., 2014). This increased importance of the bank has coincided with the recognition that more comprehensive management approaches including ecosystem-based management (ESBM) and MSP are needed to ensure maximization of opportunities and minimization of threats posed by each activity on the economic, social and environmental sustainability of the bank (Baldwin *et al*., 2014).

The biggest threats to the ecology of the bank are activities such as fishing which may lead to habitat change or destruction. Habitat destruction has been recognized as the greatest threat to biodiversity worldwide by the Millennium Ecosystem Assessment board (2005). A number of ecologically and commercially important species which contribute to the lives and ivelihoods of many Jamaicans stand to be negatively affected (Fisheries Division, 2014). Of these species, Queen conch is perhaps the most important as the stock on the Pedro Bank is perhaps the largest and most viable remaining Queen conch stock in the region (CRFM, 2012).

Part of the success of Queen conch on the Pedro Bank is due to the banks' benthic structure which provides ideal habitat for the various benthic life stages of the species which tend to follow the distribution pattern of their main macroalgal and detritus food as well as the flat sandy areas used as spawning grounds (Stoner *et al*., 1996b). The banks' benthos generally consist of various species of macroalgae and sand flats which make up two-thirds of the substratum along with a number of patch reefs and shoals interspersed with seagrasses, gorgonians, sponges and macroalgae (Dolan, 1972). The distribution of the different substrate types are not uniform across the bank as substrate structure and complexity appear to change with depth (Table 1) (Smikle and Appeldoorn, 2003). As a result, for the purposes of Queen Conch stock assessments, the bank has historically been classified into four depth zones in order to capture the relationship among; depth, substrate and conch distribution (density or abundance) (Fisheries Division, 2013).

Table 1. Summary of benthic profile for the three main depth zones used in the management of Queen conch on the Pedro Bank. Adapted from Dolan (1972) and Smikle and Appeldoorn (2003).

Depth (m)	Main habitat type	Estimated area (km²)
0-10	shallow reefs, shoals and sand substrate with irregular profile	438
11-20	Large flat sand plains with patch reefs having a more regular profile	2,338
21-30	Sand blanket comprised of carbonate, biogenic and sand detritus and macroalgae	3,346

The shallowest depth zone down to 10 m covers an area roughly 438 km² along the southeastern section of the bank. It has an irregular substrate profile due to a relatively

well-developed reef system with a number of shoals intermixed with reef pavement (weathered reef surface), coral rubble and sand flats (Dolan 1972; Smikle and Appeldoorn, 2003). This depth zone is known to have the highest density of Queen conch (averaging well over 100 conch/ha) as well as the highest level of fishing effort on the bank (Smikle and Appeldoorn, 2003; Fisheries Division, 2013). The next depth zone from 10 to 20 m, extends around 2,338 km² from the southwest and central areas through to the eastern end of the bank. Here there exists a more flat and regular substratum primarily comprised of large sandflats interspersed with patch reefs, and algal patches (Munro and Thompson, 1983). Densities of conch have been reported to fluctuate based on the results of previous abundance surveys and experiences low to intermediate levels of fishing effort (Fisheries Division, 2013).

The final zone assessed for management purposes is down to 30 m depth which is approximately 3,346 km² in area and has more extensive sand plains with deep coral rubble, shellfish and other biotic fragments weathered and transported by currents, along with various marcoalgae species (Dolan, 1972; Smikle and Appeldoorn, 2003). This is effectively the deepest zone in terms of conch scientific research and fishing activity mainly due to safe diving considerations therefore fishing activity, and indeed conch stock assessments, on the Pedro Bank do not go beyond this depth (Fisheries Division, 2013). Areas on the bank deeper than 30 m form a thin area around the periphery consisting of sandy substrate down to depths of around 70 m, the most extensive of which is located on the western side of the bank (Figure 2), which then rapidly falls away to the oceanic abyss (Smikle and Appeldoorn, 2003).

3 Materials and methods

3.1 Data

3.1.1 Description of current distribution

A survey of the Queen conch population on the Pedro Bank was carried out during the month of October 2015 to (i) provide a description of the current distribution and popualtion structure and (ii) provide data for models estimating the species' realized niche. For the latter, the 2015 survey data will be combined with data from the previous three surveys conducted in the years 2002, 2007 and 2011.

The survey involved underwater sampling at 80 sites spread across the bank down to a maximum depth of 30 m using a stratified systematic sampling design (Ehrhardt and Valle-Esquivel, 2008) (Figure 3). This design was geared toward maximizing coverage of the benthic substrata and benthic biological range occupied by the species as well as other important environmental factors such as depth differences and known fishing grounds.
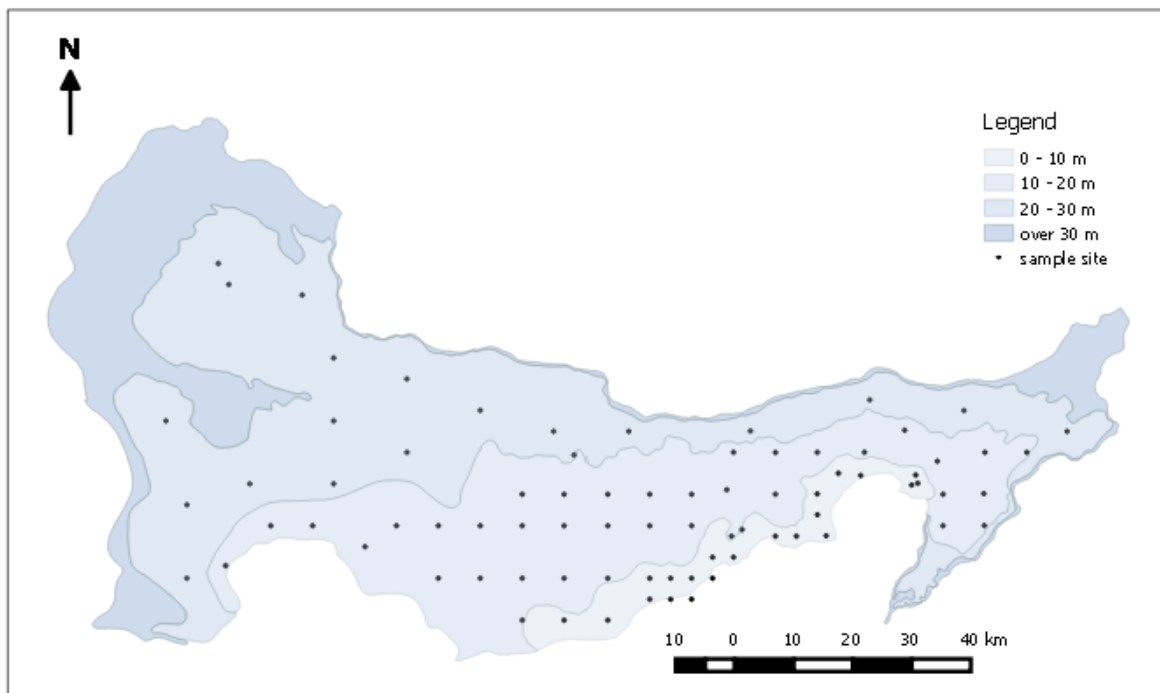
Figure 3. Stratified systematic sampling design of the 2015 Pedro Bank Queen conch survey. Stratification is based primarily on depth and substrate complexity. Map produced using QGIS (QGIS Development Team (2015).

Based on this design, the bank was stratified into three depth zones at 10 m intervals down to the 30 m depth contour where safe diving was possible. The shallowest zone, down to 10 metres, was densely sampled with 21 sites in order to cover as much of its complex benthos as possible. This zone consists of much of the traditional fishing grounds and is also known to have high abundance of conch (Fisheries Division, 2013). The next two depth zones are larger in area but have progressively less complex substrate with depth and are known to have less fishing effort for conch (Smikle and Appeldoorn, 2003). These were sampled with 40 and 19 sites respectively.

Each site in the survey was sampled, and therefore is defined, by a transect of dimension 30 m by 12 m in addition to two replicate transects placed roughly one metre from the end of the last transect. In other words, each site consisted of three transects representing a sample area of 30 m x12 m x 3 = 1,080 m². Transects were surveyed by two scientific divers at a time, one for each half. In order to reduce bias, experienced divers were used and training was conducted to ensure that subjectivity of interpretation was minimized. Recordings were made underwater of the substratum type, transect depth as well as counts of different size classes of Queen conch based on a classification scheme modified from Appeldoorn (1988) (Table 2).

Table 2. Description of Queen conch size classification. Adapted from Appeldoorn (1988).

Category	Description
Small juvenile	< 150 mm shell length
Medium juvenile	151-200 mm shell length
Large juvenile	> 200 mm shell length, but without flared shell lip, <3 years old
Sub-adult	Flared lip starting to grow, but not fully developed (lip < 4 mm thick), ~ 3 years old
Adult	Flared lip is fully formed, with minimal to moderate shell erosion, > 3 years old
Stoned conch	Shell characterized by heavy to serious erosion and heavy fouling (coral, sponges, algae, etc.). Shell lip thick and worn, > 5 years old

These size classes are determined by a combination of estimated biological criteria and ontogenetic milestones such as shell length (maximum length along the shell's horizontal axis) at sexual maturity which is around 200 mm and 3 years old. The structure of the flared lip (thickened, curved edge of shell) is also an important characteristic in determining size classes and is among the most common used criteria to estimate the age of

the species (Ehrhardt and Valle-Esquivel, 2008). Such a multi-criteria classification is useful and necessary for gastropod species such as Queen conch whose accurate ageing is often hindered by processes such as bio-erosion of the calcareous shell as well as differing growth rates at different times during its life history (Stoner and Sandt, 1992).

3.1.2 Distribution modelling

The descriptive and predictive modelling processes were carried out using the statistical and geographic information system (GIS) platforms available in the R software (R Core Team, 2014) and following modelling procedures as outlined in Zuur *et al*., (2009, 2010). Biological and environmental data from a number of sources were incorporated in the SDMs (Table 3).

Table 3. Summary of environmental and biological data used for species distribution modelling.

Data type	Data layer	Source/Reference	Purpose/Comments
Biological	Conch counts	Current study; Smikle and Appeldoorn (2003); Fisheries Division (2014).	Underwater point survey observations at sites across the Pedro Bank.
Environmental	Depth (discrete)	Current study; Smikle and Appeldoorn (2003); Fisheries Division (2014).	Obtained from field survey observations. For use in describing current distribution and estimating realized niche.
	Depth (gridded)	IOC, IHO, BODC (2003).	Gridded bathymetry data used for depth raster layer in predictive models.
	Substrate	Baldwin *et al*., (2014).	Broad scale habitat map of the Pedro Bank by Schill.
	Chlorophyll-a concentration	GES-DISC/NASA, (2014).	Monthly MODIS 4 km resolution gridded remotely sensed data. Preliminary data used.

3.1.2.1 Species biological data

For distribution modelling purposes the abundance (counts) of each size class, transect depth per site and the site coordinates from the October 2015 survey were added to similar data compiled from the previous three conch surveys conducted in 2002, 2007 and 2011. The data from these surveys were obtained from the Jamaican Fisheries Division of the Ministry of Agriculture and Fisheries database (Fisheries Division, 2014). The scope of sampling for each survey vary somewhat in terms of coverage, the number of sites sampled and the time of year that each was conducted (Table 4). In total the combined dataset included 266 sites across the Pedro bank.

Table 4. Summary of conch abundance survey design and coverage for the years 2002, 2007, 2011 and 2015 (Smikle and Appeldoorn, 2003; Fisheries Division, 2013; Fisheries Division, 2014).

Year	Survey month	Survey design	No. of sites	Transect area (m²)	Total area sampled (m²)
2002	May and December	Random with replication of transects	45	90	25,560
2007	January and November	Random with replication of transects	60	450	60,000
2011	November to December	Stratified with replication of transects	81	360	100,440
2015	October	Stratified with replication of transects	80	360	86,400

The six size classes recorded during the survey were further consolidated into three groups instead for use in distribution modelling. The three groups include; (i) "Juveniles" consisting of small, medium and large juveniles, (ii) "Mature" consisting of the main sexually mature groups; sub-adults, adult and stoned conch size classes, and (iii) "Total conch". The purpose here was to model each group separately and compare the distributions.

Counts of each group were used as the main proxy for conch abundance and thus were used as response variables in both descriptive and predictive distribution models. The use of counts instead of conch densities, which are commonly reported for conch surveys throughout the Caribbean region (Erhardt and Valle-Esquivel, 2008), was to allow for the use of more flexible models such as GAM and GAMM which can incorporate, or be extended to incorporate, more suitable data distribution and variance structures than linear regression and is also able to maintain relative abundance within each group (Reiss et al., 2014).

3.1.2.2   Environmental data

A broad scale habitat map of the Pedro Bank developed by Schill during the Pedro Bank Marine Spatial Planning Project (MSP) (Baldwin *et al.*, 2014) was used as the substrate layer for both descriptive and predictive models. The habitat layer included seven levels (habitat categories) describing or summarizing the main substratum types occurring on the bank. These include; (i) coral reef, (ii) deep coral, (iii) deep ocean, (iv) land, (v) marcoalgal hardground, (vi) sand and sediment, and (vii) seagrass. The R programming software (R Core Team, 2014) "raster" (Hijmans, 2015) and "sp" packages (Bivand *et al.*, 2013) were used to create the substrate raster layer from which information corresponding to each site coordinate were extracted. During this process diver-mounted video footage taken at

randomly selected sites during the 2015 survey were used for cross checking information from the habitat layer. This was done as the reliability of substrate data as a predictor variable in distribution modelling largely depends on ground-truth sampling (Reiss et al., 2014).

The substrate raster, discrete transect depth data and the year of each survey were used as explanatory variables in describing the species' distribution. For the predictive modelling aspect, the substrate layer was reused along with depth (gridded bathymetry data) and chlorophyll-a concentration (mg/m³) which was used as a proxy for primary production especially with regards to the distribution of marcoalgae. These particular variables were chosen because they were relatively easy to access along with the fact that they are among the primary environmental factors known to influence Queen conch distribution (Stoner and Lally, 1994). The bathymetric depth layer was developed from the General Bathymetric Chart of the Oceans (GEBCO) Digital Atlas published by the British Oceanographic Data Centre (BODC) on behalf of the Intergovernmental Oceanographic Commission (IOC) and the International Hydrographic Organisation (IOC, IHO, BODC, 2003) while the chlorophyll-a concentration layer was developed based on MODIS 4 km resolution ocean data retrieved through the Goddard Earth Sciences Data and Information Services Center's (GES-DISC) Interactive Online Visualization and Analysis Infrastructure (Giovanni) GES-DISC/NASA, (2014).

3.2 Data exploration

Data collected during the October 2015 survey and the combined surveys datasets were put through a series of exploratory procedures in order to decipher the most salient patterns. For the former the aim was simply to describe the structure of the data, but for the latter data exploration was aimed at understanding the data structure with a view toward guiding the distribution modelling process. A thorough and guided data exploration process is essential to modelling ecological data as data taken directly from nature often do not conform to assumptions of traditional modelling tools such as the linear regression model (Zuur *et al.*, 2010). Data exploration here was therefore geared toward exploring and testing for possible violation of the linear regression model assumptions which include; (i) normality in the distribution of covariates, (ii) homogeneity of variance in the response variables, and (iii) independence (non-correlation) between covariates (Zuur *et al.*, 2009).

A number of visualization tools available in R were used to carry out the data exploration process. The pairs function from the R package "graphics" whose output include scatterplots, histograms and pair-wise correlation measurements (R Core Team, 2014) were applied to both datasets. These plots were used to examine the data for normality, the relationship between response and explanatory variables, as well as possible dependencies between and among covariates. To check for heteroscedasticity or heterogeneity of variance the boxplot, also from the "graphics" package, was used which gives an idea of the spread of response variable (counts) among different nominal variables (size classes, survey year, and substrate types). Conditional plots, or coplots, from the "lattice" package (Sarkar, 2008) were also used to examine for possible two-way interactions between response and explanatory variables conditioned on one or more of the nominal variables.

3.3 Current distribution (2015 survey)

3.3.1 Density and abundance

The 2015 survey data was used to describe the current Queen conch distribution on the Pedro Bank by calculating densities estimates for each site and also for each depth zone. The procedure involved first taking an average of the three transects (one transect and two replicates) at each site then calculate the density per zone area for each different group/size classes. Abundance could then be calculated for each zone by multiplying the densities by the respective area of interest. This can be summarized by the formula:

$$Abundance = \left(\frac{Number\ of\ conch}{Area\ of\ transect}\right) * Area\ of\ stratum$$

Given the biology and expected distribution of the species especially with respect to its tendency to occur in patches across its range, it may be reasonably assumed that their distribution (density and abundance) would not be expected to follow a normal distribution nor have homogeneity of variances (Ehrhardt and Valle-Esquivel, 2008). Also it is quite common in ecology to have over-dispersion or the situation where variance is larger than the mean (Zuur *et al.*, 2009). Therefore, in order to obtain reasonable 95% confidence intervals for the density and abundance estimates means of each were resampled using the bootstrap method (Efron and Tibshirani, 1993). The procedure was implemented with the functions boot and boot.ci in the R "boot" package (Davison and Hinkley, 1997; Canty and Ripley, 2014).

3.3.2 Spatial distribution

The calculated densities for each size class were plotted over-laying the habitat layer in order to visualize their spatial distribution with respect to the different habitat types. This was conducted using the functions from the R packages "sp" (Bivand *et al.*, 2013), "raster" (Hijmans, 2015), "GISTools" (Brunsdon and Chen, 2014) and "maps" (Becker *et al.*, 2014).

3.4 Distribution modelling

3.4.1 Realized niche

3.4.1.1 Model fitting and selection: GAM

Guided by the results of the data exploration, the dataset of combined survey data was fitted to an appropriate model in order to describe the species' realized niche. This was done assuming that no major explanatory variable was omitted from the data fitted to the model. A general additive model (GAM) was chosen as an appropriate starting model as its very robust algorithms are appropriate for count data and are also able to deal with data such as this where violation of linear regression assumptions are likely (Hastie and Tibshirani, 1990). Linear models were not chosen due to their relatively poor performance in adequately modelling ecological data as a result of their rather rigid assumptions which

often do not fit data taken directly from nature (Guisan and Zimmermann, 2000; Zuur *et al*., 2009). GAMs on the other hand allow for several non-linear relationships between response and explanatory variables and can deal with heterogeneity of variances as well as dependent response values through the inclusion of spatial and temporal correlation structures (Zuur *et al*., 2009).

The general form of the GAM model applied to the data is given by the formula:

$$Y_i = \alpha + f(X_i) + \varepsilon_i \qquad \text{where } \varepsilon_i \sim N(0, \sigma^2)$$

where *Yi* is the response variable, *α* is intercept parameter, and *f* (*Xi*) is the smoothing function for the smoother. ε_i is the residuals containing the unexplained information of the model defined as the difference between observed and fitted values (Zuur *et al*., 2009). Residuals are assumed to be normally distributed with a mean of *0* and variance *σ²*. This basic model can be further extended to yield hybrid multivariate forms where explanatory variables may be continuous or nominal as is the case in this analysis.

The specific model chosen was a GAM assuming Poisson distribution of the response variable *Y* and an offset variable (intensity parameter) to address bias due to differences in transect area (Zuur *et al*., 2009). The Poisson distribution is ideal for count data having lots of zeroes, and confers the advantage of the probability for negative values being zero (number of conch cannot be less than zero) and also allowing for heterogeneity in the mean variance relationship (Guisan *et al*., 2002; Zuur *et al*., 2009). The full model is summarized by the formula:

$$\log(Y_i) = \alpha + f(X_{1i}) + f(X_{1i}) * factor(Z_i) + factor(W_i) + \log(V_i) + \varepsilon_i$$

Here the term $\log(V_i)$ is the natural log of the offset variable incorporated to address heterogeneity issues due to the differences in the area surveyed at each site during different years. The full model with variables inserted is given by:

$$\log(count_i) = \alpha + f(depth_i) + factor(substrate_i) + f(depth_i) * factor(substrate_i)$$
$$+ factor(year_i) + \log(area_i) + \varepsilon_i$$

Here $\log(count_i)$ is the number of conch (assumed to be Poisson distributed), $f(depth_i)$ is the smoothing function for transect depth at each site, $substrate_i$ and $year_i$ are used as factors within the model having six and four levels respectively. Also included in the GAM were an interaction term between the depth smoother and substrate type as well as a term for the residuals given by ε_i.

Since this GAM is intended to be a starting model it was run for only the "Total" conch group and would only be run for the two subgroups if the model validation was favorable. The gam function in the R package "mgcv" by Wood (2004) was used the run the model and details of the model were obtained using the summary and anova commands (R Core Team, 2014).

3.4.1.2 Model validation: GAM

Validation of the GAM was conducted by examination of the output of the gam.check function from the "mgcv" package which includes; a residual Q-Q plot, a histogram of the residuals, a plot of the response versus fitted values, and plot of the residuals versus linear predictor (Wood, 2004). These were used to check for normality, homogeneity and overall model fit. Spatial correlation (dependence) was also tested for especially since failure to identify and deal with correlation in spatial data can potentially invalidate important statistical tests such as the F-test (Zuur *et al.*, 2009). A residual variogram was used, which is a useful tool for measuring spatial dependence between sites where sites that are closer together given a specific metric are more similar than those further apart (Zuur *et al.*, 2009). The bubbleplot was also used which is a plot of the residuals versus their spatial coordinates where clusters of negative or positive residuals are indicative of dependence (Zuur *et al.*, 2010). Both plots were implemented in R using the functions variogram and bubble from the "gstat" package (Pebesma, 2004).

3.4.1.3 Extending the GAM: GAMM

Given the nature of the data (data from different surveys over a spatial range) the results of the validation exercise of the GAM may be expected to reveal evidence heteroscedasticity and spatial correlation in the residuals. The GAM was therefore extended to a generalized additive mixed model (GAMM) to improve model fit through the addition of optimal variance and spatial correlation structures. An important aspect of this procedure was to first determine the optimal variance and correlation structures for the three conch groups: total, mature and juvenile conch.

The optimal variance structure for the data was obtained by using the general least squares (GLS) algorithm as given in the "nlme" package in R (Pinheiro *et al.*, 2014). Using the total conch and the covariates the GLS model was fitted with different variance structures based on Pinheiro and Bates (2000). There are several variance structures which are better for different types of data. Those selected for testing here were those best for data with likely heterogeneity along the levels of nominal variance covariates (substrate, year) where different variances are expected for each factor level (Zuur *et al.*, 2009). They are also able to allow for increasing or decreasing variance along a continuous variance covariate such as depth in this case. The variance structures added to the model included; (i) the constant variance function structure implemented in R with the function varIdent, (ii) power variance function (varPower), (iii) the constant plus power variance (varConstPower), (iv) exponential variance (varExp) structure, and finally (v) a combination of variance functions (varComb). The best correlation structure was selected using the Akaike Information Criteria (AIC) (Akaike, 1973) implemented in R using the AIC function from the "stats" package (R Core Team, 2014). This optimal variance structure was then added to the GAMM for each conch group.

The optimal spatial correlation structure was obtained by firstly adding different structures to the GAMMs of each conch group. Correlation structures included; exponential, linear, Gaussian, Rational quadratic and spherical correlation structures based on Schabenberger and Pierce (2002) and implemented using the gamm function in the package "mgcv" (Wood, 2004). The through a process of backward selection using AIC the best correlation structure was selected. An alternative to this approach to using AIC would have been to extend the model to include additional covariates and their possible interaction which

could have better explained the response (Zuur *et al*., 2009) however this option was outside the scope of this project as additional covariates was not an option.

3.4.1.4   Model validation: GAMM

Validation of the optimal GAMM included examination of the residuals of the GLS models with the optimal variance structures for each conch group. Here a plot of the normalized residuals versus fitted values were used along with a plot of normalized residuals versus the explanatory variable depth (the only continuous covariate). The second aspect was to check for spatial dependence using variograms for the different conch groups. Both were implemented using functions from the "nlme" and "graphics" R packages (Pinheiro *et al*., 2014; R Core Team, 2014).

3.4.2 Predictive modelling

3.4.2.1   GAM prediction

The GAM component of the optimal GAMM describing the ecological profile for total, mature and juvenile conch were extracted and used to produce a habitat map of predicted abundance for the entire Pedro Bank including unsampled areas. All variables were included except that depth referring to transect depth at each site was replaced by the gridded bathymetry data extracted from the General Bathymetric Chart of the Oceans (GEBCO) Digital Atlas published by the British Oceanographic Data Centre (BODC) on behalf of the Intergovernmental Oceanographic Commission (IOC) and the International Hydrographic Organisation (IOC, IHO, BODC, 2003). Predicted values were produced using the predict.gam function from the "mgcv" package then converted to a raster object with functions from the "raster" package (Hijmans, 2015).

3.4.2.2   Maxent prediction

A predictive habitat map estimating the habitat suitability for Queen conch on the Pedro Bank was also produced using the Maxent software (Phillips *et al*., 2006). The programme was accessed and implemented in R through functions available in the "dismo" package (Hijmans *et al*., 2015). Maxent essentially estimate the probability of species presence by modelling species presence based on the given environmental factors and a random (or specified) background sample of the study area (Phillips *et al*., 2006; Elith *et al*., 2011).

From the total of 266 sample sites in the combined surveys dataset 233 had total conch values greater than zero and were therefore used as presence values for the model. Of these presence values a portion of 20% was withheld to be used as test points for evaluating the model, while the remaining points were used as training data for model fitting. The background data was defined arbitrarily as 250 random points falling within a 4 km radius around each training point (Figure 4). This was done using the circles and spsample functions from the packages "dismo" (Hijmans *et al*., 2015) and "sp" (Bivand *et al*., 2013).
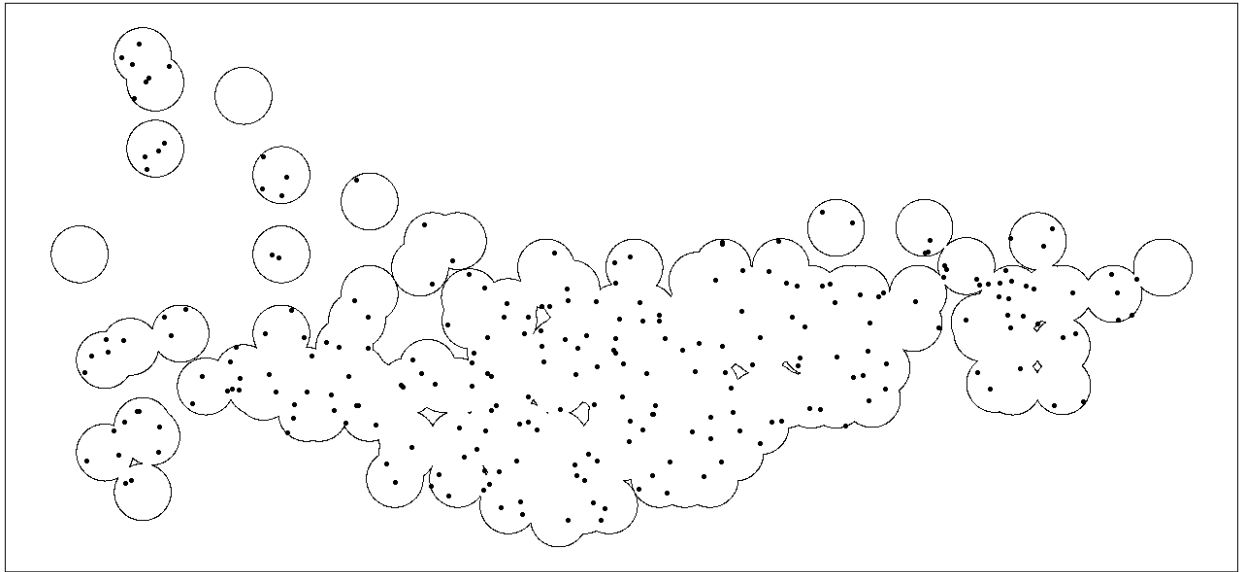
Figure 4. Illustration of the relative position and coverage of random background points used in the fitting of the Maxent model. Circles represent a 4 km radius around each sample point. Adapted from Yoder (2013).

Restricting the area in which background points would fall was done in order to have the random background points falling as much in the study area as possible and in reasonable proximity to presence sites and not in deeper areas that were not sampled and were unlikely to have the species present (Elith *et al*., 2011; Yoder, 2013).The background sample is important as it informs the model about the density of covariates across the entire sample area thus providing the basis for comparison with the density of covariates occupied by the species (Elith *et al*., 2011). The model was then run in R with default settings and the salient model features produced and examined using the maxent, response, predict, evaluate and plot R functions.

Model evaluation involved a k-fold cross-validation procedure utilizing the 20% withheld portion of the presence points (Phillips *et al*., 2006). The evaluate function (Hijmans *et al*., 2015) was the main tool used to implement this in R. The function produces an object containing a number of model evaluation metrics, including; true positive rate (TPR) and false positive rate (FPR) derived from confussion matrices, the models' associated AUC (Area under a receiver operating characteristic (ROC) curve) value and an association measure (Hijmans *et al*., 2015; Phillips *et al*., 2006). These are robust metrics which are among the most commonly used for evaluating Maxent models and other logistic-type models (Raven, 2002). AUC is threshold independent and is able to measure predictive accuracy or the probability that a randomly chosen presence location is ranked higher than a randomly chosen background point (Merow *et al*., 2013).

4 Results

4.1 2015 abundance survey

4.1.1 Data exploration

The first aspect of data exploration was focused on the 2015 conch survey data which included records from the six size classes. The first exploratory tool applied here was a pairplot of the number or counts of each size class and transect depth (Figure 5). The results showed evidence of non-normal distribution among the size classes. In fact, the data distribution closely resembles a Poisson distribtion which is typical of count data extracted from nature (Zuur *et al*., 2009). In this case, sites of low counts occurred disproportionately more frequently than sites with higher counts. The distribution of transect depths within the sample was however normally distributed. The pairplot also indicated no clear relationship or interaction between transect depth and the numbers of each size class. Among the size classes there were some indications of pair-wise interaction particularly between size classes close to each other in terms of length for example between medium and large juveniles. Interaction among other size class pairs however seemed to be very little for instance bewteen small juveniles and adults.
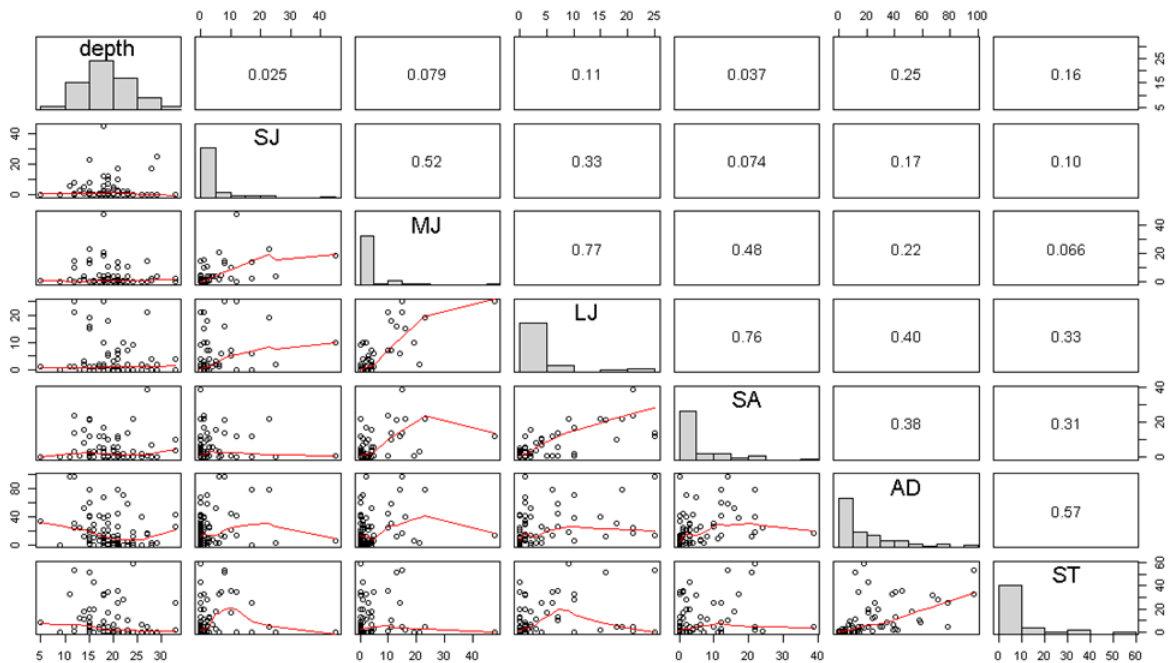
Figure 5. Pairplot of the total number of conch sampled within each size class along with transect depth from the 2015 conch survey dataset. The respective upper and lower panels relative to the diagonal show pair-wise Pearson correlations and scatterplots. Coding for

the size classes are as follows: SJ= small juvenile, MJ = medium juvenile, LJ = large juvenile, SA = subadult, AD = adult and ST =adult stoned conch.

Boxplots of the number of conch per size class showed that there were different data spreads among them and also a fair number relatively extreme values compared to the median (Figure 6). This would be somewhat expected given the species' tendency to cluster and aggregate in response to various ecological and environmental effects (Stoner and Lally, 1994). The plot results are also consistent with a highly skewed, non-normal data distribution. Median values are similar across the size classes except for adults and to a lesser extent stoned conch which have higher values and a much larger data spread.
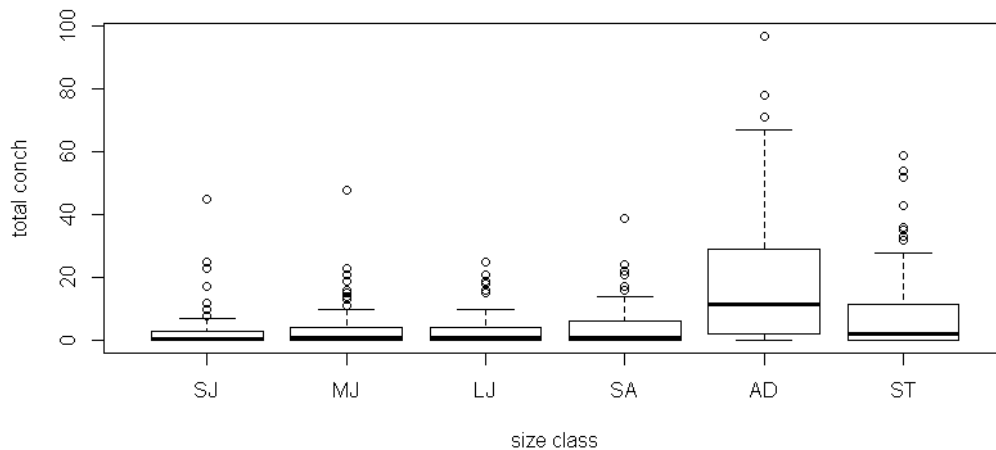
Figure 6. Boxplot of the total number of conch sampled within each size class from the 2015 conch survey. The coding for the size classes are as follows: SJ= small juvenile, MJ = medium juvenile, LJ = large juvenile, SA = subadult, AD = adult and ST = adult stoned conch.

A boxplot was also produced of the number of conch among the three depth zones. This indicated that there were different data spreads among the depth zones as well as a much more distinct difference in median values (Figure 7). The 0-10 m zone had the largest median and data spread while the 21-30 m zone had the lowest median and data spread.
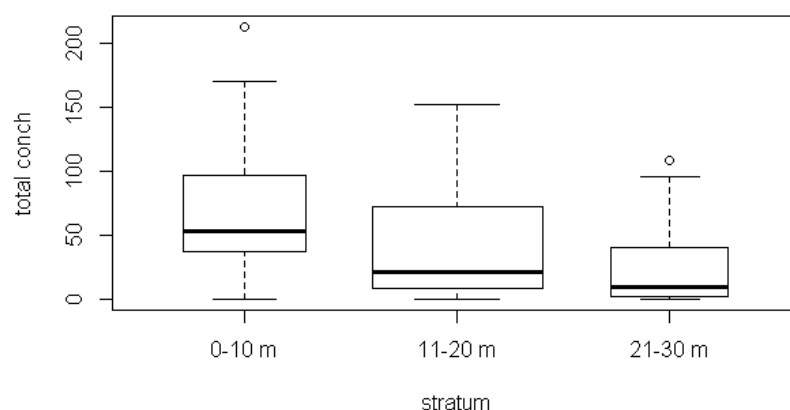
Figure 7. Boxplot of total number of conch sampled at sites within the three main depth zones/strata from the 2015 conch abundance survey.

4.1.2 Density estimates

The overall total density of conch on the bank was 409 conch/ha while the juveniles and mature conch groups had mean densities of 98 and 311 conch/ha respectively (Table 5). The mature conch density figure is significant as it is well above the minimum 56 reproductive adult conch/ha reference point suggested by Stoner and Ray-Culp (2000) below which *Queen conch* populations may begin to experience negative population growth due to the allee effect.

Table 5. Total conch density (conch/ha) and bootstrap obtained 95% confidence limits for each size class across the Pedro Bank from the 2015 survey.

	Small juvenile	Medium juvenile	Large juvenile	Subadult	Adult	Stoned conch	Total density	Mature conch	Juveniles
Density (hectare)	30	36	33	44	182	85	409	311	98
Lower confidence limit	15	19	20	27	133	56	318	236	60
Upper confidence limit	43	49	44	59	228	112	504	383	130

The two most important reproductive size classes, adults and stoned conch, were in fact individually above this reference point while subadults and the other juvenile size classes had relatively similar in mean densities (Figure 8).
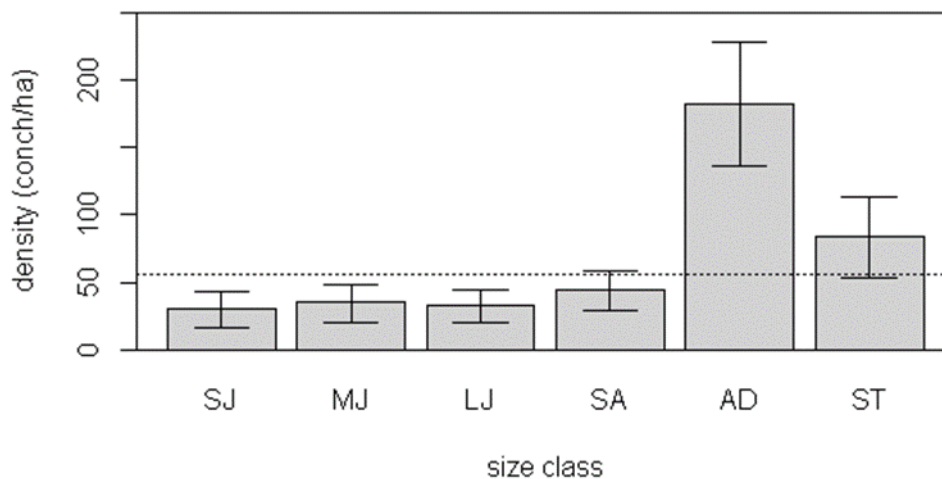
Figure 8. Barplot of density (conch/ha) with 95% bootstrap confidence limits of each size class. The dotted horizonal line refer to the 56 adult conch per hectare reference point suggested by Stoner and Ray-Culp (2000) below which Queen conch populations may begin to experience negative population growth. The coding for the size classes are as

follows: SJ= small juvenile, MJ = medium juvenile, LJ = large juvenile, SA = subadult, AD = adult and ST = adult stoned conch.

Densities of size classes within each depth zone had a more or less similar distribution to the overall density on the bank (Figure 9). The important reproductive size classes (adults and stoned conch) had densities that were above the minimum 56 conch/ha reference point except in the 10-30 m zone. Densities of juvenile size classes in each zone were relatively similar to each other.
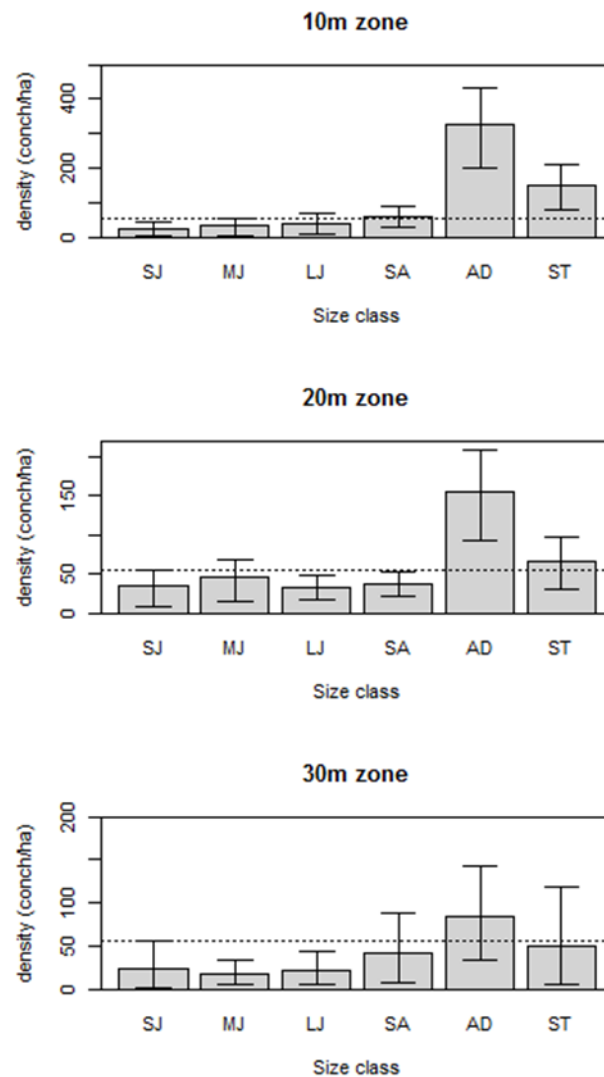
Figure 9. Barplot of density (conch/ha) with 95% bootstrap confidence limits of size classes within each surveyed depth stratum. The dotted horizonal line refer to the 56 adult conch/ha reference point suggested by Stoner and Ray-Culp (2000) below which Queen conch populations may begin to experience negative population growth. The code for the size classes are as follows: SJ= small juvenile, MJ = medium juvenile, LJ = large juvenile, SA = subadult, AD = adult and ST = adult stoned conch.

4.1.3 Spatial distribution of conch densities

In terms of their spatial distribution, the highest total conch densities occured at sites located at the southeastern section of the bank, particularly in the shallowest depth zone (0-10 m) where the substrate consists of mostly sand sediment, seagrass, coral reefs, and small areas of marcoalgal hardground (Figure 10). Total conch density were relatively lower in the other areas of the bank, though the western side had a few sites with high densities well over 300 conch/ha. This area was not heavily sampled due to its irregular depth profile, strong current and issues affecting diver safety.
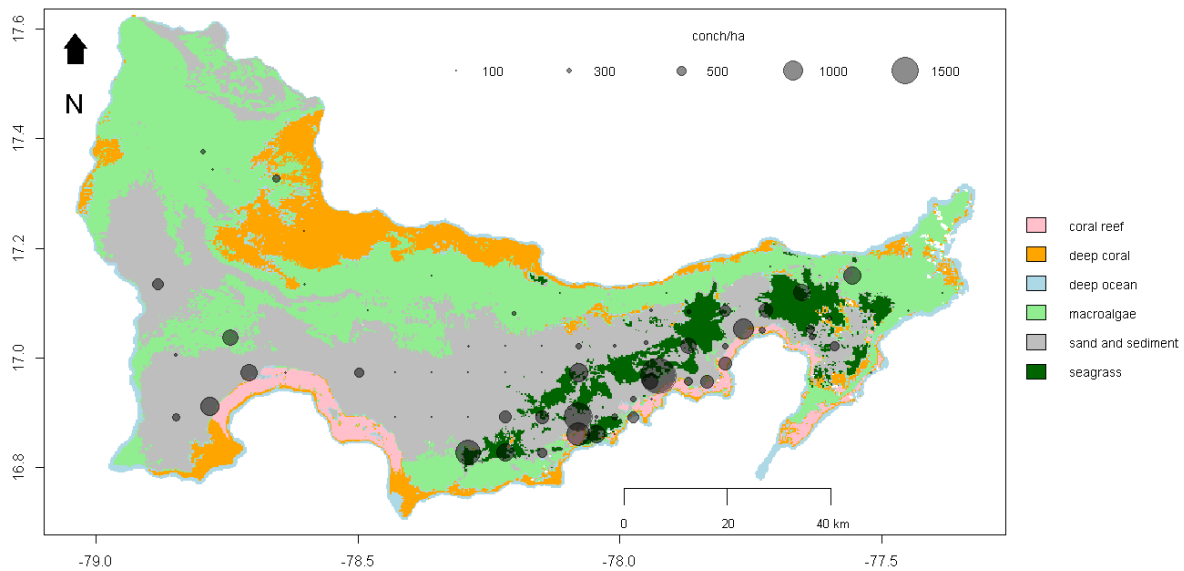
Figure 10. Broad scale habitat map of the Pedro Bank (Baldwin et al., 2014) overlaid by a bubbleplot of total conch density at each 2015 survey site.

Similar plots of site densities for mature and juvenile conch. They showed that sites of highest denisties for both occured generally in the same areas and among the same set of substrate types. The distribution of mature conch density followed closely the pattern of total density suggesting that the mature portion of the population is accounting for most of the total conch density observed on the bank (Figure 11). Density of mature conch throughout other areas of the bank was fairly even except for a few high density sites at the western end of the bank.
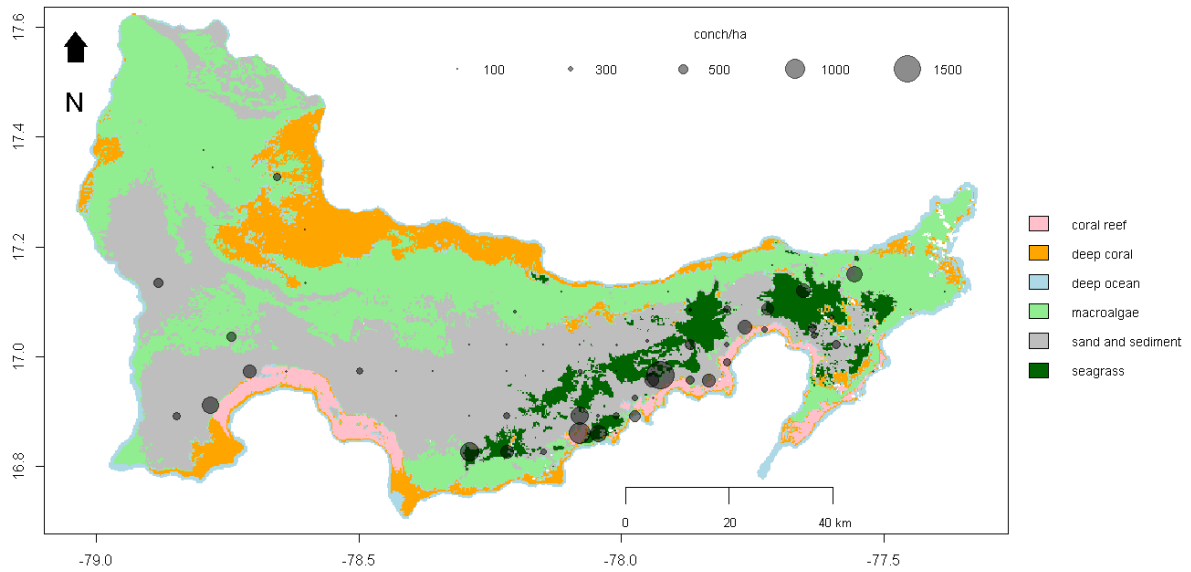
Figure 11. Broad scale habitat map of the Pedro Bank (Baldwin et al., 2014) overlaid by a bubbleplot of mature conch density at each 2015 survey site.

Juvenile densities was also fairly even among sites across the bank except for small clusters of highest density sites at the southeast of the bank associated with macroalgae, sand and seagrass (Figure 12). There were also a few isolated high density juvenile sites that stood out in the extreme western and eastern areas as well.
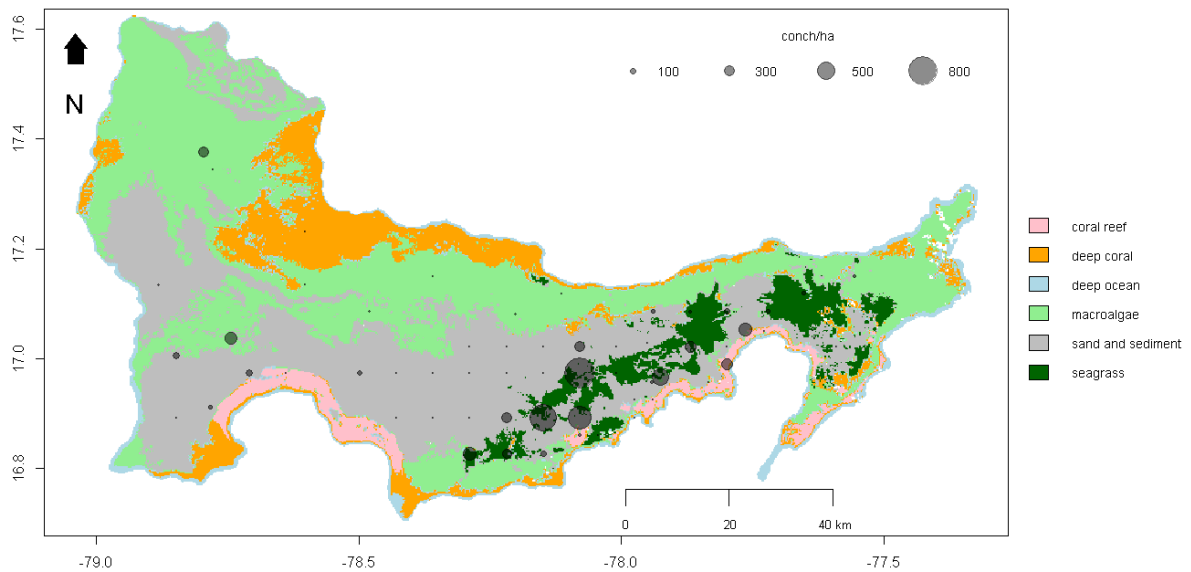
Figure 12. Broad scale habitat map of the Pedro Bank (Baldwin et al., 2014) overlaid by a bubbleplot of juvenile conch density at each 2015 surveyed site.

4.2 Distribution modelling

4.2.1 Data exploration

The second data exploration exercise was applied to the combined survey dataset obtained from the 2002, 2007, 2011 conch survey as well as the 2015 survey. The goal here was to guide the distribution modelling process in terms of fitting the most appropriate models based on the structure of the data in order to estimate the species' realized niche and make predictions based on the findings. As previously mentioned, for modelling purposes the size classes were grouped into three categories; namely, (i) "Juveniles" consisting of small, medium and large juveniles, (ii) "Adults" consisting of the main sexually mature groups; sub-adults, adult and stoned conch size classes, and (iii) "Total" conch which includes all conch specimens.

A pairsplot was made of the counts of the size groups along with the variable transect depth to investigate their data distribution and pair-wise relationships (Figure 13). There was very little indication of interaction or dependence between the groups and depth, neither were there clear evidence of interaction or dependence between the juvenile and mature conch groups. There appeared however to be a high level of interaction between the juvenile and mature groups and total number of conch indicated by their pair-wise correlations and scatterplots. This is expected since both subgroups are nested within the total variable. The histograms within the pairsplot also showed a that depth is normally distributed and that the counts of each conch group appeared to be Poisson distributed where there is a high proportion of low values.
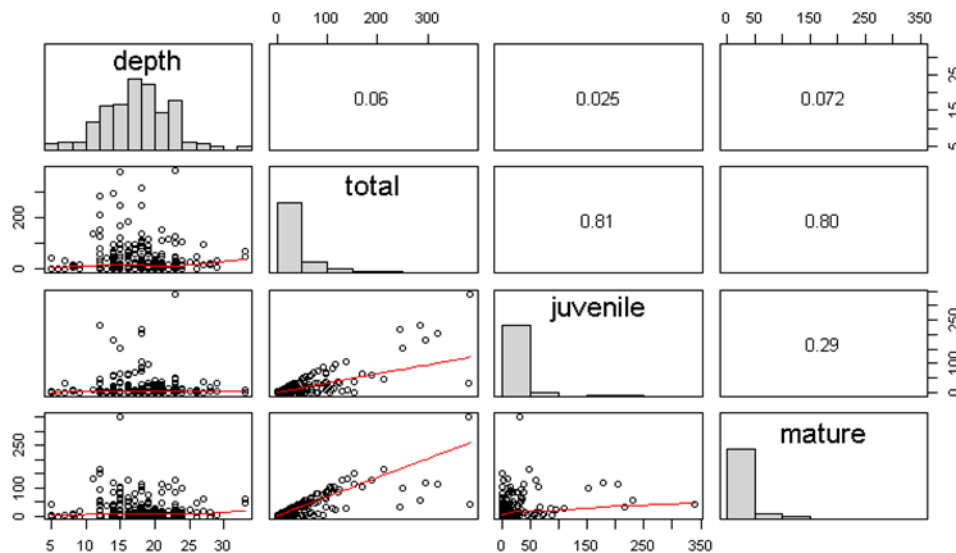
Figure 13. Pairplot of transect depth, total conch, juvenile conch and mature conch. The center panels show histograms of each variable while the left and right panels respectively show pair-wise relationship scatterplots and Pearson correlations.

The data spread for total conch among the four survey years and among the various substrate types both revealed evidence of heteroscidasticity among the two categorical vaiables (Figures 14 and 15). In both cases there appeared to be a different data spread for

each level of the two factors. Median values were fairly similar in both cases as well and most values occurred well above the the median which suggests a skewed non-normal distribution of the data and potential for influential values if a linear regression was to be applied to the data (Zuur *et al*., 2010).
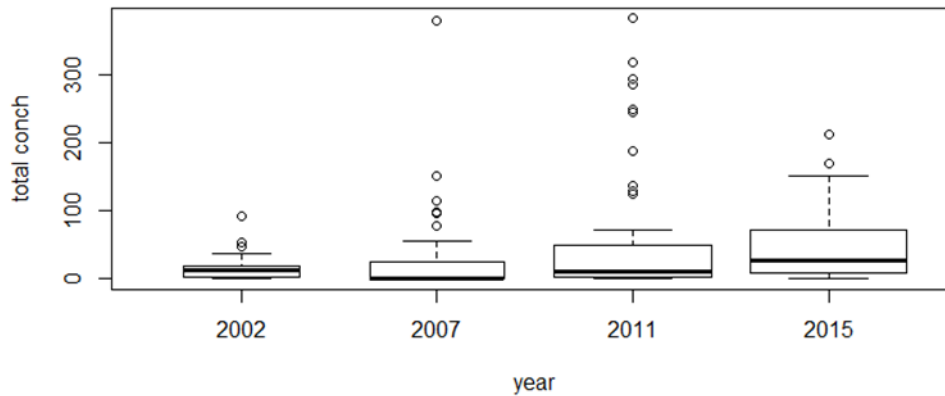
Figure 14. Boxplot of total number of conch sampled at each site for each survey year. There is eveidence of heterogeneity among the survey years.
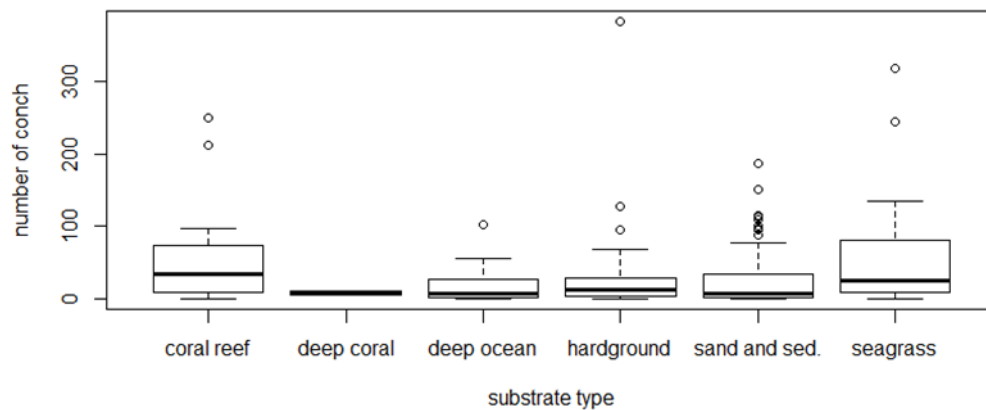
Figure 15. Boxplot of the total number of conch among substrate types for the combined surveys dataset. There is evidence of heterogeneity among the subtrate types as well as likely skewed non-normal distribution in the data.

The next step in the data exploration exercise was to further investigate two-way and possible three-way interactions among the variables. In other words, explore whether the number of conch at different depths change with regards to substrate type among the survey years. This was done using a conditioning plot or coplot from the R "lattice" package (Sarkar, 2008) (Figure 16).
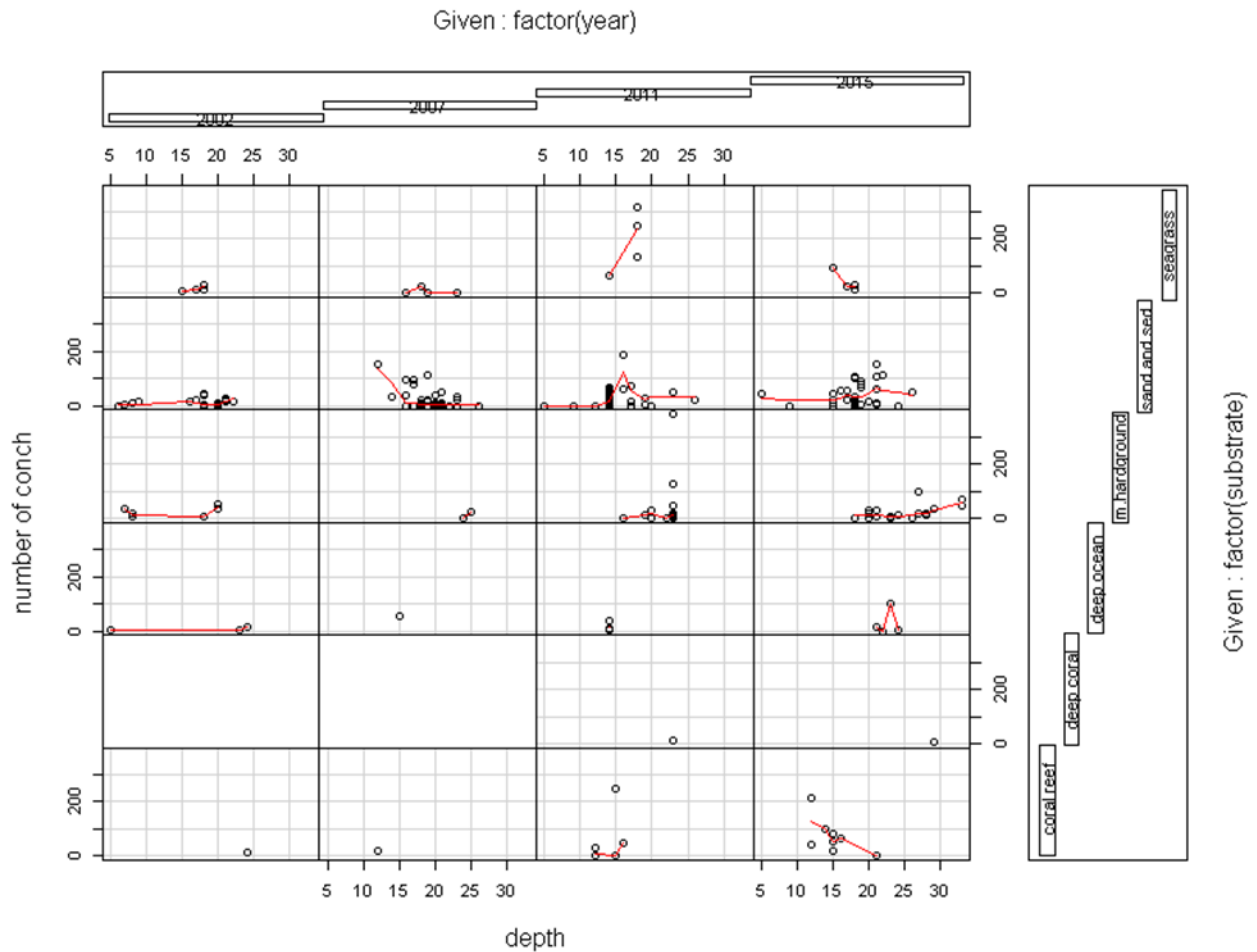
Figure 16. Coplot of combined conch survey dataset including scatterplots of the total number of conch versus depth conditioned on the survey year and substrate type as well as added smoothers to aid visual interpretation.

The resulting coplot with added smoother showed that there may be a three-way relationship among the variables. The relationship is however not very clear in all cases as there are different numbers of sites for each survey year and substrate type. For instance, there is a high number of sites on macroalgal hardground and sandy substrate, and also more overall number of sites in 2011 and 2015 compared to the other two survey years. Notwithstanding this there is still reason enough to consider this possible interaction in the distribution modelling process. The depth variable appear to only cause small changes in total conch although in most cases this is not entirely clear as there are only a few sites occuring for certain years and substrate types.

4.3 Estimating the realized niche

4.3.1 Starting model: GAM

A GAM assuming Poisson distribution as previous described was fitted to the combined surveys dataset to estimate the realized niche of the Queen conch on the Pedro Bank. The model included total conch (counts) as the response variable and the explanatory variables;

depth (smoother), the depth-substrate interaction (smoother) as well as the terms substrate and year which are nominal variables. The output summary of the model as produced by the R anova function (R Core Team, 2014) revealed that all terms were highly significant (Table 6). The anova function was used instead of other summay functions avaialable in R because it uses the F-statistic which is best for testing the significance of nominal variables with more than two levels as is in this case (Zuur *et al*., 2009).

Table 6. Summary output of the starting GAM given by the anova R function.

Term type	Variable	Effective degree of freedom	Degree of freedom	Chi-square	p-value
smoother	Depth	8.979	-	373.8	<2e-16
smoother	Depth:substrate	9.446	-	991.6	<2e-16
parametric	Substrate	-	4	248.4	<2e-16
parametric	Survey year	-	3	4986.4	<2e-16

These results would suggest that counts of total conch change among substrate types and among survey years, and also that the overall effect of depth and substrate is highly influential in the distribution of conch across the bank (Figure 17).
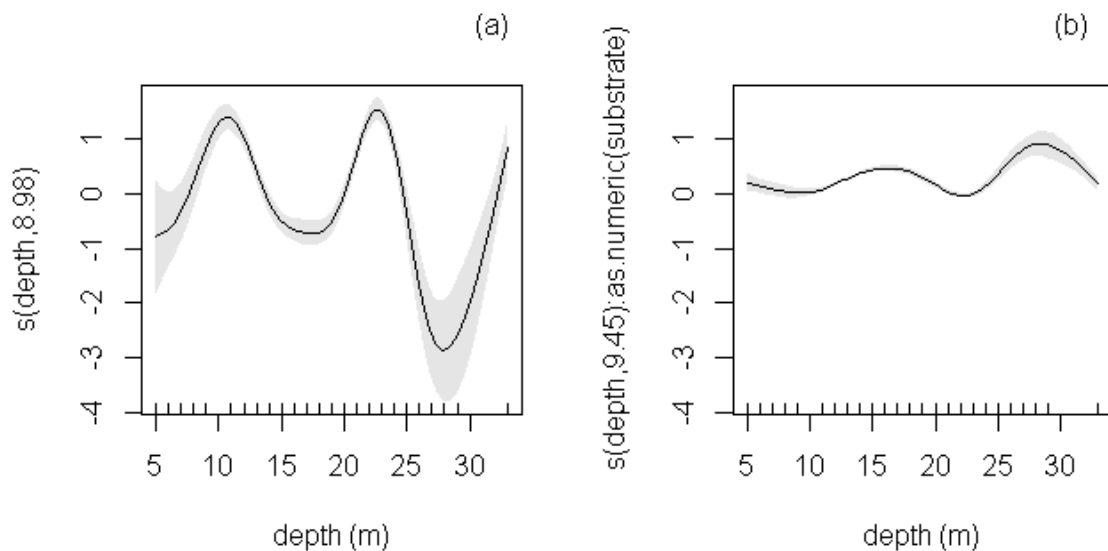
Figure 17. Estimated GAM smoothing curves for depth (a) and depth-substrate interaction (b).

4.3.2 Model validation: GAM

Validation of the starting GAM was carried out through the examination of its residuals which revealed that the model results were not reliable and could be further improved (Figure 18). Plots of the residual distribution showed evidence of heteroscidasticity as there is a clear pattern where the residual spread appears to get larger with respect to increasing values of the linear predictor and fitted values. There is also evidence of non-normality in the distribution of the residuals though this maybe a less severe violation since regression models can be robust against some amounts of non-normality (Fitzmaurice *et al*., 2004). In this case however the departure from normality is severe.
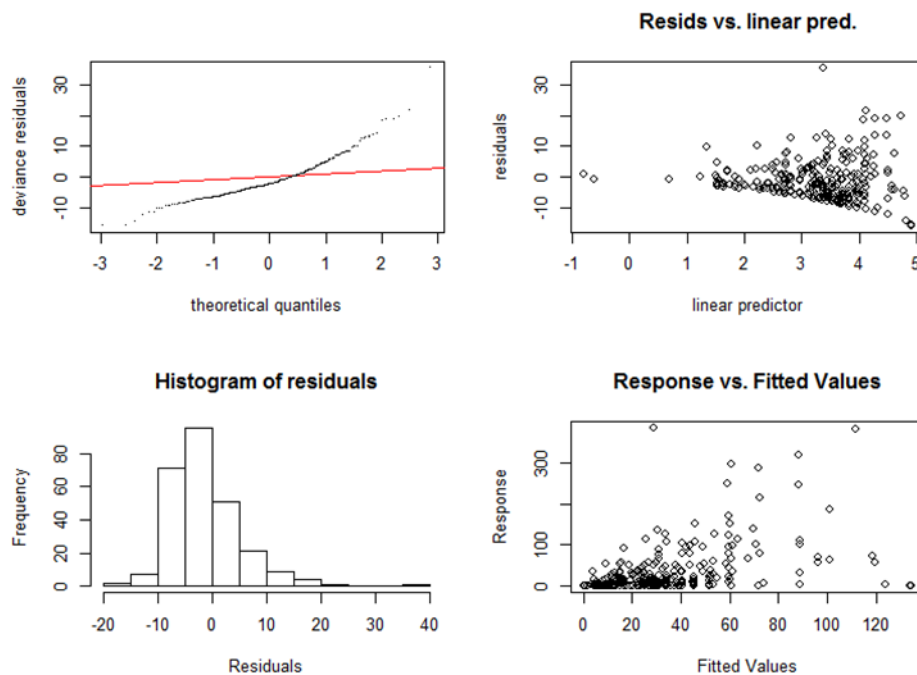
Figure 18. Residual plots for fitted starting GAM as given by the gam.check R function. There are patterns in the residuals as well as increased residual spread from low to high linear predictors and fitted values which is a sign of heteroscedasticity.

Further examination of the residuals was carried out to check for spatial independence among values at each sample site across the bank. This was done firstly through a bubbleplot of the model residuals versus site coordinates which allows for a visual examination for possibly spatially auto-correlated sites which show up as clusters of negative and positive residual values (Figure 19) (Zuur *et al*., 2009). The plot showed evidence of likely spatial dependence among the sites.
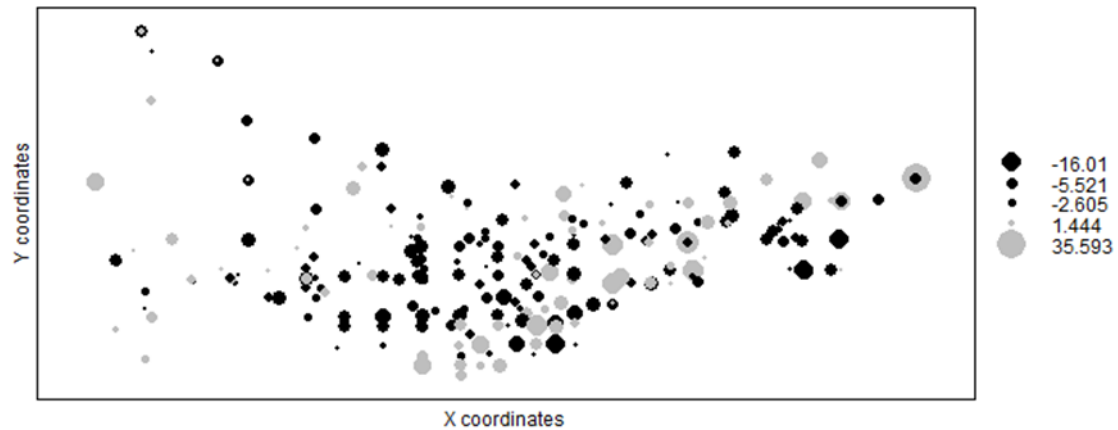
Figure 19. Spatial bubbleplot of the GAM residuals versus their coordinates. Grey circles represent positive residuals and black circles represent negative residuals, each are proportional to the size of the residuals. Clusters of positive and negative residual values indicate likely spatial dependence.

Secondly, a residual variogram which is a more formal spatial dependence meaure was produced using the varigram function from the "gstat" R package (Pebesma, 2004). Variograms are a more robust tool measuring spatial dependence between sites based on distances between their spatial coordinates (Zuur *et al*., 2009). The values produced in the variogram suggests that there is spatial denpendence (Figure 20) as there is an increase over a range of the lower distance values roughly from 0 to 30,000 which then levels off at higher distances.
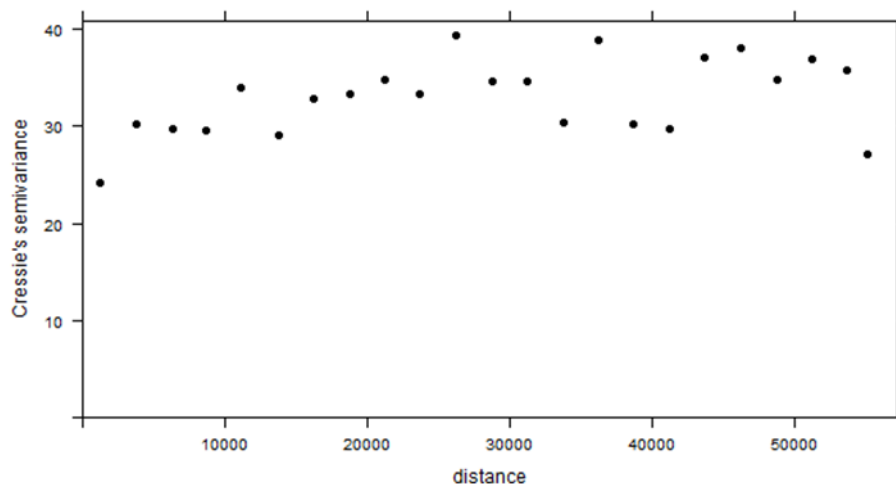
Figure 20. Cressie's semivariance rubust variogram of residual values from the GAM. Values appear to increase from a nugget over a range of lower distance values, roughly between 0 and 30,000, then come to an asymptote afterwards. This is indicative of spatial correlations.

These validation results for the GAM dictates that the model must be further improved to address, in particular, the issues of spatial dependence among sites and heterogeneity as a result of different survey years and different substrate types.

4.3.3 Extending the GAM: GAMM

4.3.3.1   Choosing the best corelation structure

Given the model validation results, the GAM was extended to a general additive mixed model (GAMM) to incorpoate spatial correalation and variance structures in order to deal with spatial dependence and heterogeneity in the data and improve model fit. The first step was to choose the best correlation structure for the data. A number of correlation structures were fitted to the model including; exponential, linear, Gaussian, Rational quadratic and spherical correlation structures (Schabenberger and Pierce, 2002). The best model for each of the groups; total conch, mature and juveniles conch, were chosen through a process of backward selection based on a comparison of AIC values (Table 7).

Table 7. Akaike Information Criteria (AIC) values for different correlation structures applied to the GAMM for each conch group. AIC values were used in the backward model selection process to choose the best correalation structure for each conch group. In all cases AIC values showed significant improvement compared to the starting GAM (AIC=12,975).

Correlation structure	Degree of freedom	AIC value for Mature	AIC values for Juveniles	AIC values for Total
None	-	-	-	12,975.00
Gaussian	13	1,092.32	1,180.19	1,021.22
Exponential	13	1,092.32	1,180.19	1,021.22
Linear	13	1,070.93	994.84	994.84
Rational quadratic	13	1,080.16	1,164.33	994.77
Spherical	13	1,070.27	1,180.19	994.05

The optimal correlation model for total conch and mature conch included the spherical spatial correlation structure obtained through the R function corSpher while the best model for juveniles had the linear correlation structure (Pinheiro and Bates, 2000). Models containing linear and rational quatratic (corRatio) correlation structures for total conch may have also been consisdered as their AIC vaules were very close to the optimal. The same could also be said of the linear correalation structure in the case of mature conch. Noteable from these results as well is the fact that each of the extended models with different correlation structures resulted in significant model improvement based on AIC compared to the starting GAM which had an AIC value of 12,975.

Validation of the optimal spatial correlation model for each group was done using residual variogram plots which showed a much improved plot compared to Figure 19 as there was very little indication of spatial dependence (Figure 21). Sites within each group no longer showed an obviously increasing pattern at lower distance values but instead occured more

or less as a cloud of horizontal points which is indicative of spatial independence among the points (Zuur *et al.*, 2009).
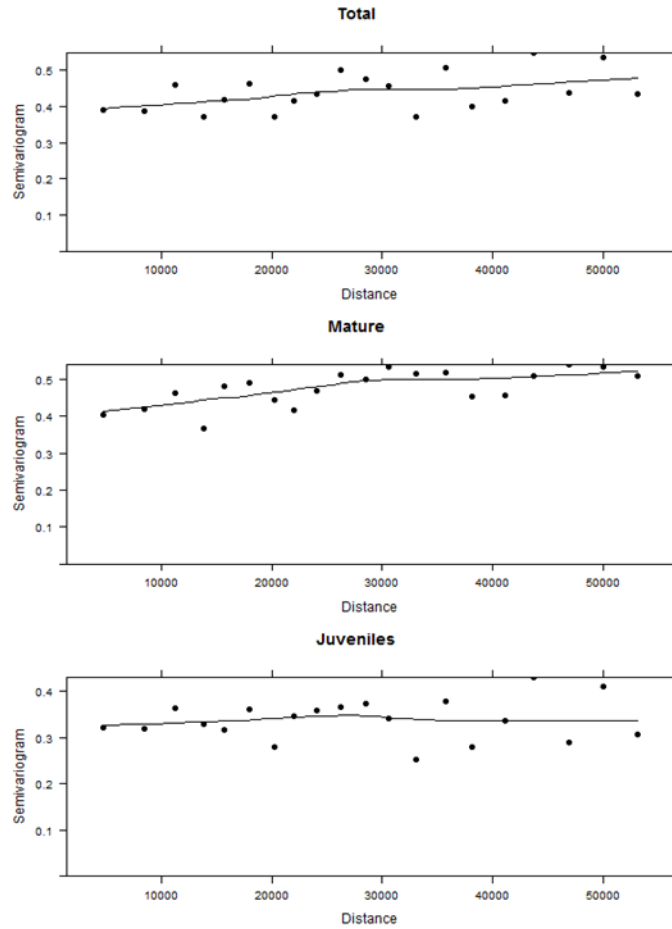
Figure 21. Residual semivariograms of the optimal GAMM for each conch group; total conch, mature and juveniles, with added loess smoother to visualize trends. The points of the variogram appear largely as an horizontal cluster across the range of distances rather than a steep increasing trend from the lower distances..

4.3.3.2 Choosing the best variance structure

The next step in extending the starting model was to determine the best variance structure to be added to the GAMM of each group to deal with heterogeneity within the data. A number of variance structures were fitted to the variable total conch using general least squares (GLS) from the R "nlme" package (Pinheiro *et al.*, 2014). The choice of variance structure was based on the data structure of the covariates and their likely residual spread based on the data exploration exercise. The constant variance function correlation structure from the "nlme" package (Pinheiro and Bates, 2000) implemented in R with the function varIdent was selected as the main structure for modelling heterogeneity among the groups. This variance structure allows for different variance for each factor level. Also included were the power of the covariate variance function (varPower), the constant plus power of the variance covariate function (varConstPower) and exponential variance (varExp), as well as a combination of variance functions (varComb) (Pinheiro and Bates, 2000).

The best variance structure was then selected through backward selection using using AIC values (Table 8). The best model was the one that included a constant variance structure (varIdent) which considered different variances for each level of the nominal terms year and substrate type.

Table 8. Results of the backward selection process using AIC to determine the best variance structure for the GAMM. The best model is model 2 having a constant variance structure (varIdent) which has the lowest AIC value. Model 1 is equivalent to a linear model and is included for comparision.

Model	Variance structure	Variance covariates	Degree of freedom	AIC for total conch
1	None	-	11	2816.26
2	varIdent	Year, substrate	32	2668.19
3	varPower	depth	12	2811.10
4	varConstPower	depth	13	2813.10
5	varComb	Year, substrate, depth	20	2688.62

Validation of this optimal variance structure model was done by plotting the models' normalized versus fitted values (Figure 22). The plot showed a much improved residual spread with no clear pattern and a more even residual spread where most values occured withing the range -1 to +1 along the y-axis.
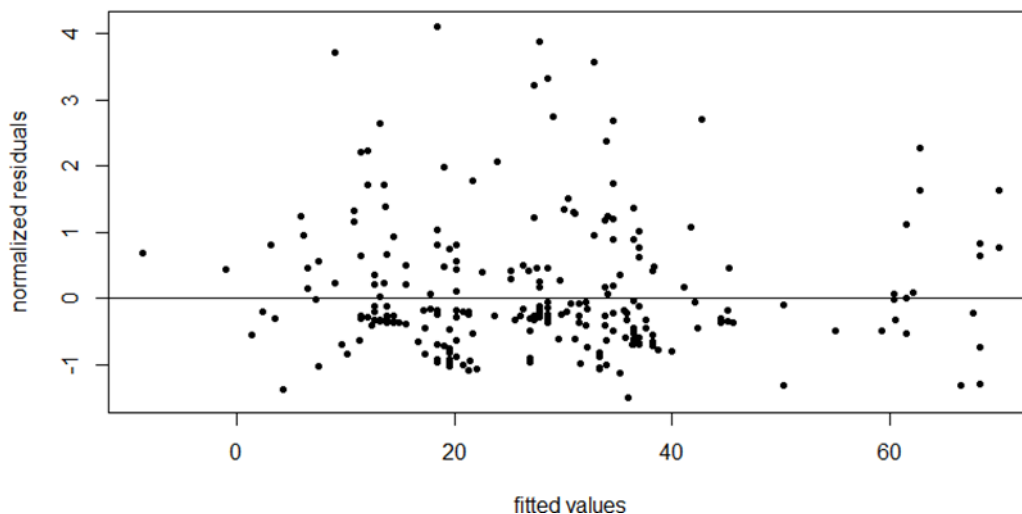
Figure 22.Plot of normalized residuals versus fitted values for the optimal correalation structure obtained through the general least square (GLS) model.

4.3.4 Optimal realized niche models

Having determined the optimal variance and correlation structures for the data these model components were then added to the final optimal GAMM for each of the conch groups. A summary as given by the R anova and summary functions for total conch revealed that all main terms including; the smoother for transect depth and both factoral terms substrate and survey year were significantsuggesting different distributions for each of the significant

model covariates (Table 9a). The model summary also showed that only 23% of the variability of the response was explained by the covariates according to the r-squared value. For the mature conch all term were significant as well (Table 9b). The depth smoother for juveniles was however found to be non-significant while the substrate and survey year terms were significant (Table 9c). Of note here as well is that the r-squared value is surprizingly negative, despite the significant terms, suggesting the model may not be fully explaining juvenile abundance.

Table 9a-c. Output model summaries for; (a) total conch, (b) mature conch, and (c) juvenile conch as given by the anova function in R.

(a)

	Total conch					
Variable	Term	df	edf	F-statistic	p-value	Adjusted R-squared
Depth	smoother	-	1.004	42.16	3.94e-10	0.226
Substrate	parametric	5	-	3.786	0.00253	
year	parametric	3	-	14.660	7.88e-09	

(b)

	Mature conch					
Variable	Term	df	edf	F-statistic	p-value	Adjusted R-squared
Depth	smoother	-	1	9.131	0.00276	0.124
Substrate	parametric	5	-	2.85	0.016	
Year	parametric	3	-	11.35	5.2e-07	

(c)

	Juvenile conch					
Variable	Term	df	edf	F-statistic	p-value	Adjusted R-squared
Depth	smoother	-	1.004	2.649	0.104	-0.00127
Substrate	parametric	5	-	4.213	0.00107	
Year	parametric	3	-	6.278	0.00040	

The smoothers for transect depth for total and mature had effective degree of freedoms (edf) of 1 (or close to 1) which means the relationship between the distribution of these groups and depth is likely to be linear in each case (Figure 23) (Zuur *et al*., 2009).
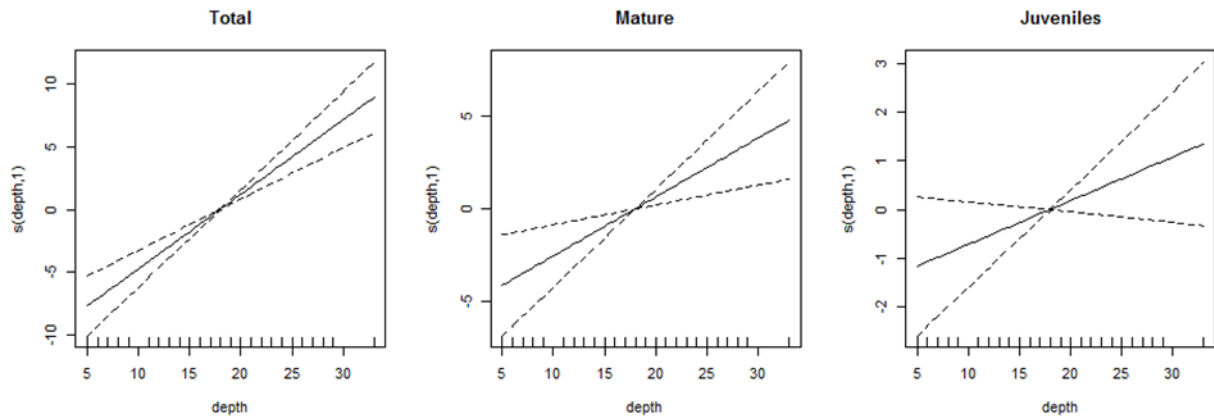
Figure 23. Plot of significant depth smoothers for total conch and mature conch and non-significant smoother for juvenile conch from the optimal GAMMs.

Regarding the different substrate types, the GAMM revealed that seagrass, sand and sediment, and marcoalgal hardground were the most important in terms of overall abundance of Queen conch on the Pedro Bank (Figure 24). Mature conch seem to favour seagrass relative to the the other substrates which had a substatially low abundances. Juveniles however, and somewhat surprisingly, had their highest abundance closely followed by seagrass.
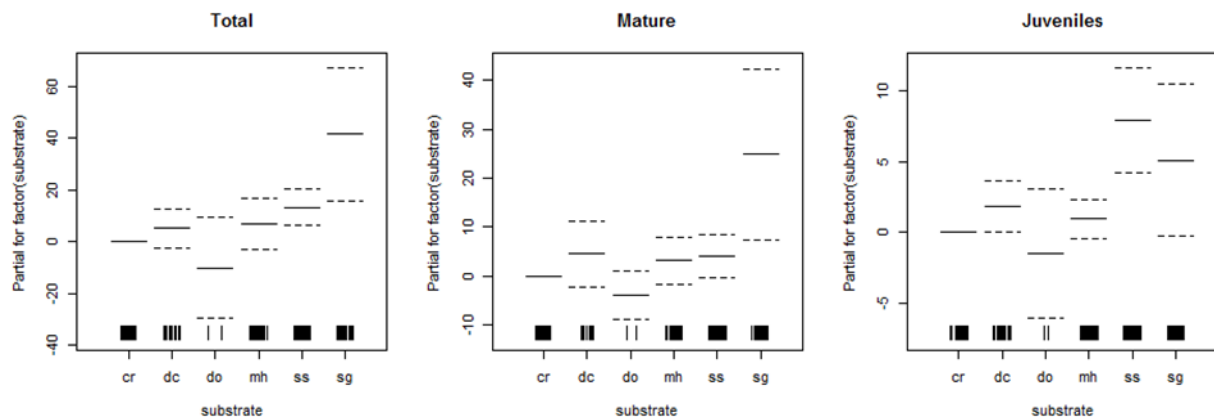
Figure 24. Plot of significant parametric substrate terms and 95% confidence limits for total conch and the mature and juvenile subsets. Total conch and the subset of mature conch have their highest abundances in seagrass while juvenile abundance was highest in sand and sediment closely followed by seagrass as well. (cr=coral reef, dc=deep coral, do=deep ocean, mh=macroalgal hardground, ss=sand and sediment, sg=seagrass).

Abundance values among the groups among the survey years were also significantly different (Figure 25). Numbers of total conch appear to steadily increase over the period of the four surveys.Mature conch seem to account for much of this increase as they showed a tremendous increase especially over the last two surveys in 2011 and 2015.
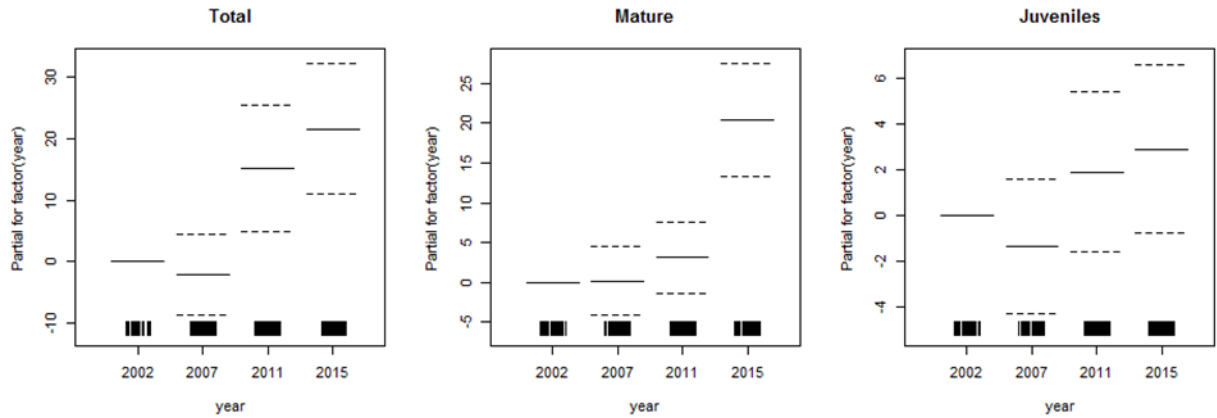
Figure 25. Plot of the significant parametric term, year, for each conch group from the optimal GAMM.

Validation of the optimal GAMM included a check for homogeneity of variance using plots of the normalized residuals versus fitted values as well as normalized residuals versus the explanatory variable transect depth for total conch (Figure 26). Since mature and juvenile conchs are subsets of the total conch it may be reasonably assumed that their residual pattern will be reflected in the residuals of the total. The residuals spread shown in each plot shows little to no residual patterning and most of the residual values fall within the range -1.5 and +1.5 therefore homogeneity can be reasonably assumed in the model.
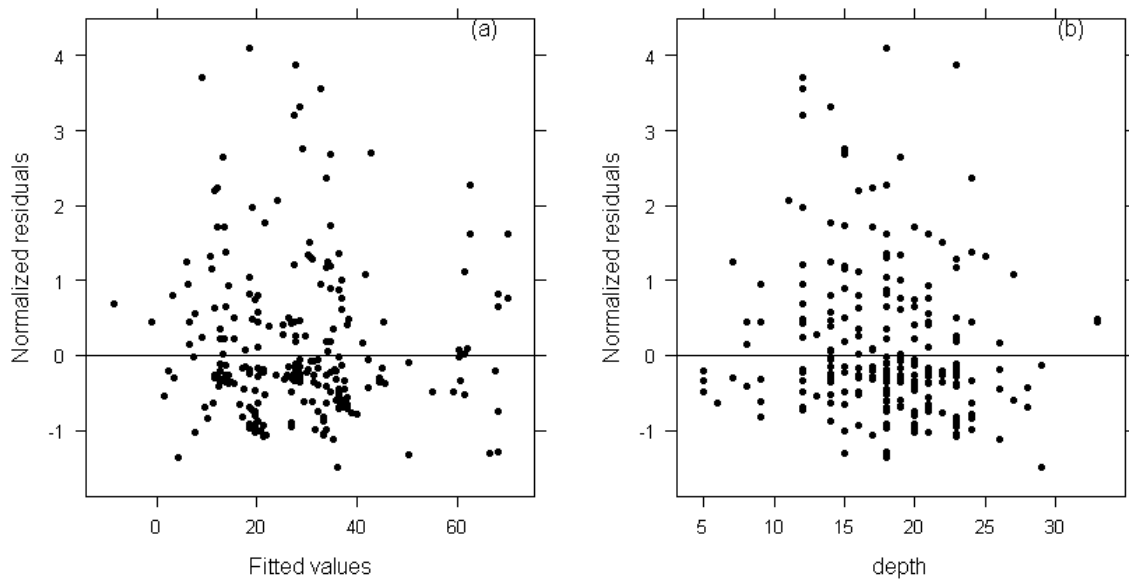
Figure 26. Residual plots of GAMM of total conch; (a) normalized residuals versus fitted values and (b) normalized residuals versus the explanatory variable transect depth. Both residual plots show very little residual patterning compared to the starting GAM.

4.4 Predictive models

4.4.1 GAM predictive model

The parameters of the optimal GAMMs estimating the realized niche were used to develop predictive models of the abundance and distribution of Queen conch on the Pedro Bank. The GAM predictions produced here are predicted probabilities of abundance and, may be cautiously interpreted as relative indices of environmental suitability (Guisan, *et al*., 2002; Phillips *et al.,* 2006). A predictive habitat map of total conch across the bank showed that there are fairly suitable conch habitat in most areas and thus the potential for relatively high abundances as well (Figure 27). Predicted abundances were highest in the southeastern region of the bank, while lowest predicted abundances occurred in those areas mainly around the periphery of the bank and in the northeast.
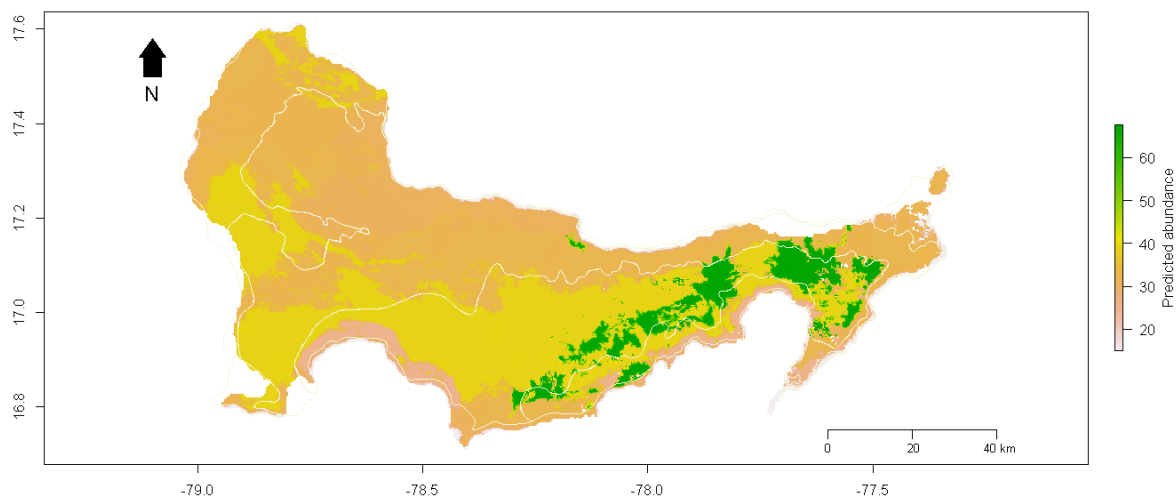
Figure 27. Habitat map of the Pedro Bank showing generalized additive mixed model (GAMM) predicted abundance of total conch.

For mature conch, the highest abundance and most likely suitable habitat corresponded largely with high abundance areas for total conch overall in the southeastern section of the bank (Figure 28). Other areas of the bank had a lower predicted abundance but were in general only slightly lower.
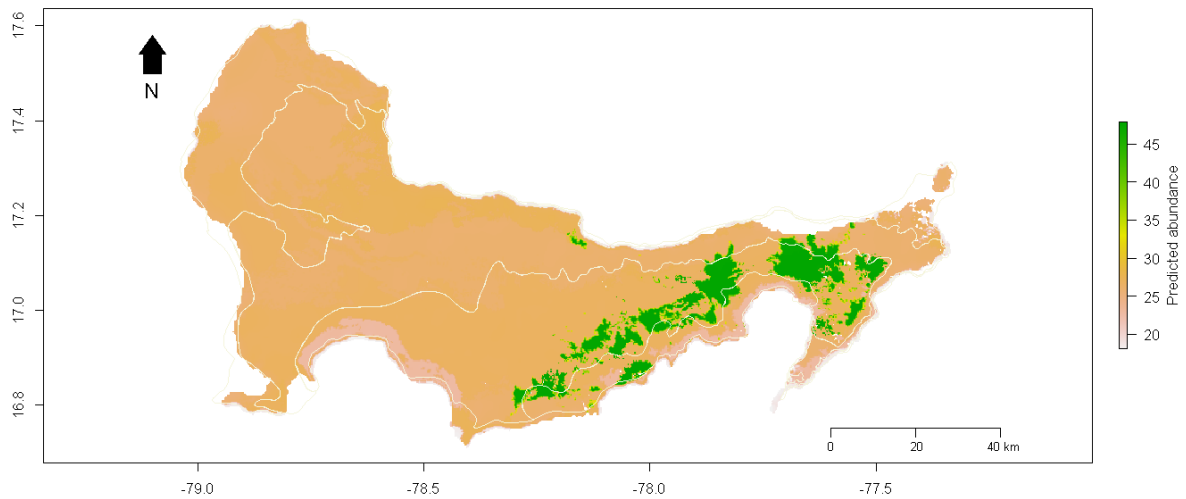
Figure 28. Habitat map of the Pedro Bank showing generalized additive mixed model (GAMM) predicted abundance of mature conch.

Habitats with highest predicted abundance and suitable conditions for juvenile conch corresponded with much of the area below the 20 m depth contour as well as a few areas to the extreme west and northwestern of the bank (Figure 29). These areas are those consisting of seagrass as well as sand and sediment.
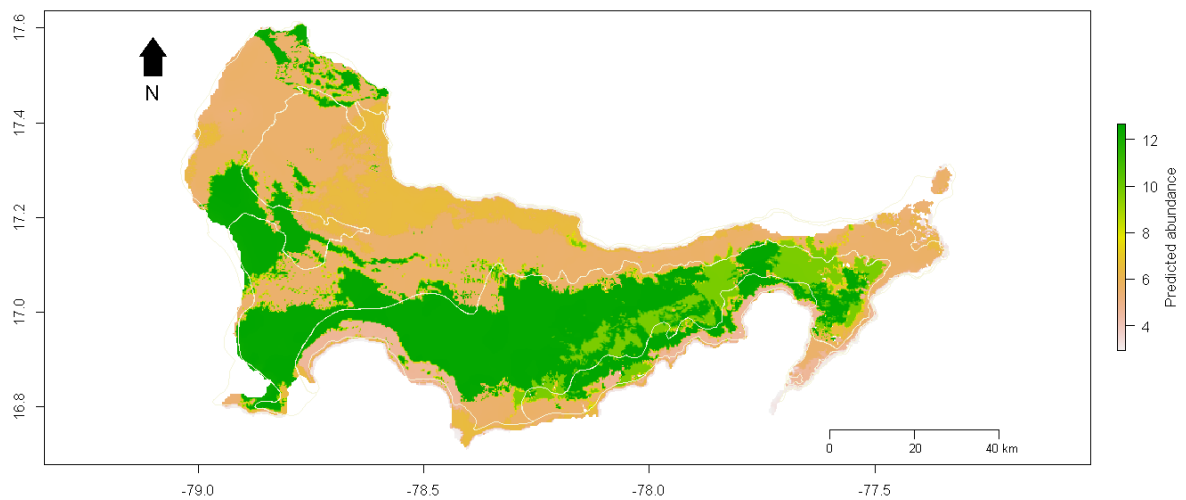
Figure 29. Habitat map of the Pedro Bank showing generalized additive mixed model (GAMM) predicted abundance of juvenile conch.

4.4.2 MaxEnt predictive model

A habitat suitability map for Queen conch on the Pedro Bank was also produced using the Maxent model. The predictions here were produced mainly to provide a comparison with the GAMM predictions and test its suitability for providing useful species distribution information specifically for this area and species. The model was based on total conch sampled (inclusive of all size classes) and revealed that the most suitable habitat lay to the southeastern and central areas of the bank, similar to the predictive GAMM (Figure 30). These areas also correspond to most areas below the 20 m contour as well where

probability of suitable habitat is largely above 0.4. The maxent model also point to a few additional areas of relatively high probability of suitable habitat which were not shown in the predictive GAMM. These include small isolated areas to the extreme east and on the south and southwestern edges of the bank. The rest of the bank to the deeper northeast section had the lowest levels of habitat suitability where probabilities were largely less than 0.5.
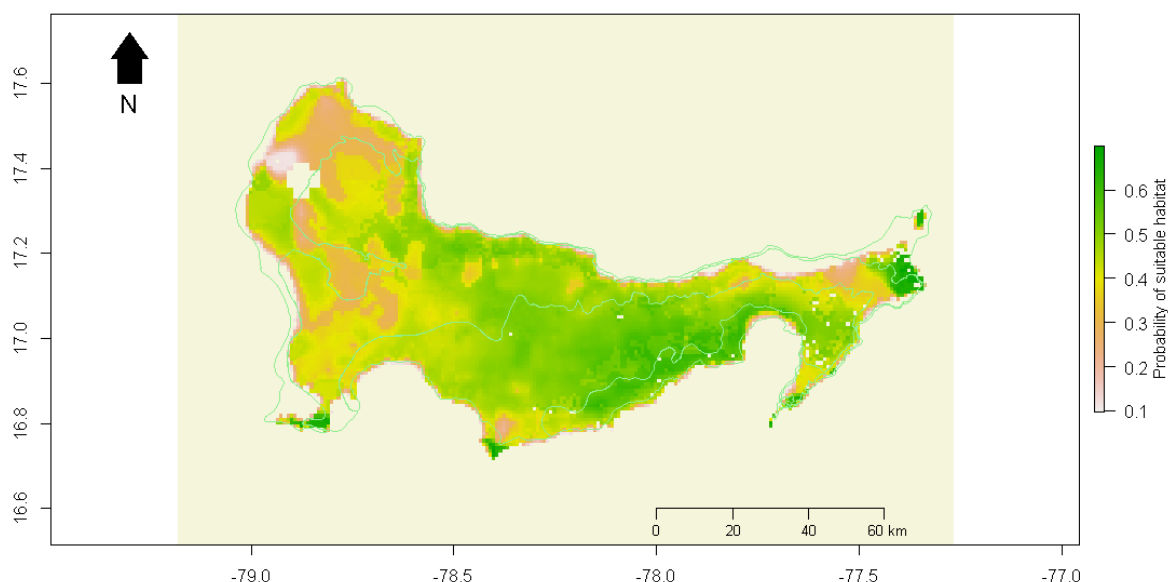
Figure 30. Map of the Pedro Bank Jamaica showing Maxent predicted probability of suitable habitat conditions.

Other output of the the Maxent prediction revealed that the variables substrate type, depth and chlorophyll-a concentration had percentage contributions of 4.66, 33.36 and 61.98 to the model prediction respectively (Figure 31). This does not necessarily indicate there importance in and ecological sense but rather their relative role in producing the Maxent model predictions.
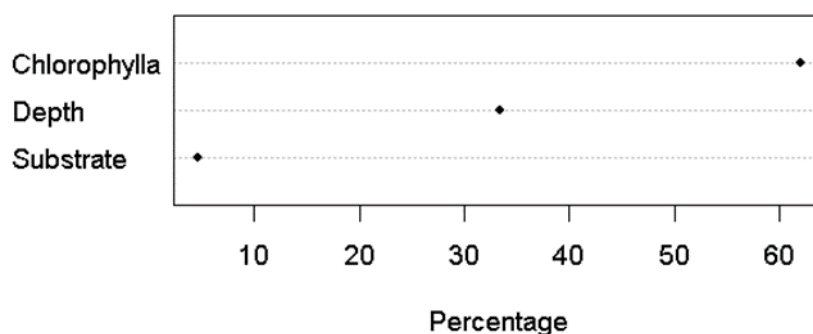
Figure 31. Precentage contribution of explanatory variables to the Maxent model prediction. The percentages given here does not necessarily reflect the ecological importance of the variables but their contribution to predictions in the maxent algorithim.

The change in predicted probabilty of suitable conditions at different values of each covariate was examined using response curves (Figure 32). These curves show how each

variable affects the prediction when all other variables are kept at their average sample value (Phillips and Dudik, 2008).
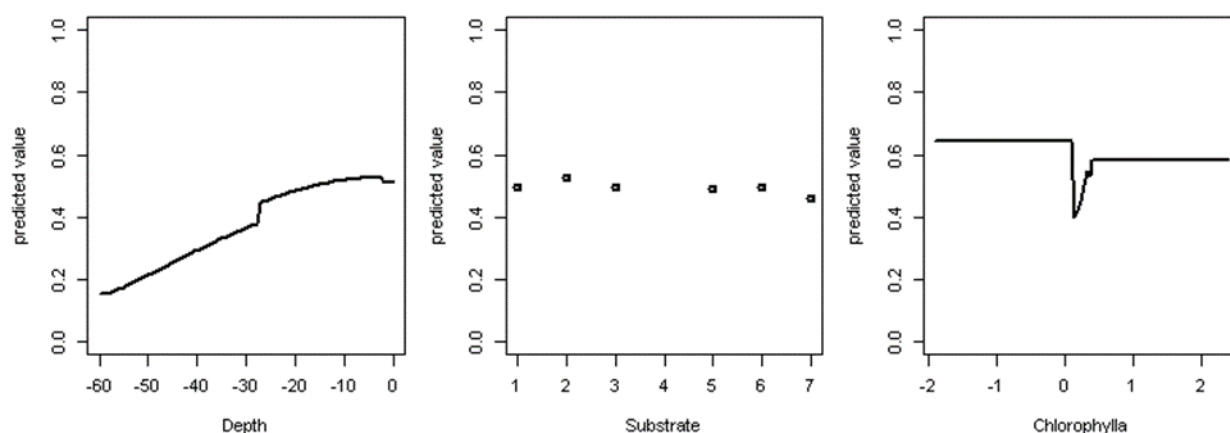
Figure 32. Response curves of the environmental variables depth (m), substrate and chlorophyll-a concentration (mg/m³). The substrate variable id coded: 1=Coral Reef, 2=Deep Coral, 3=Deep Ocean, 4=Land, 5=Macroalgal Hardground, 6=Sand and Sediment, and 7=Seagrass.

The response curve for depth showed a general decreasing probability of habitat suitabilty with increased in depth. Different substrate types also showed different responses as well with coral and deep coral areas having the highest probability of around 0.5. Other substrate types had probability values below but very close to 0.5. Positive chlorophyll-a concentrations seem to cause a sharp increasing response within short range of values then quicly reaches an asymptote.

The Maxent model predictions were evaluted using the AUC (Area under a receiver operating characteristic (ROC) curve) metric along with true positive rate (TPR) and false positive rate (FPR) derived from confussion matrices which are often used to assess model sensitivity (Raven, 2002). AUC values were computed for five different subdivisions of the data (k-fold cross validation), their values ranged between 0.61 and 0.70 (Table 10). In each case the AUC was better than random (0.5) and thus indicate that the model discriminates fairly well between suitable and unsuitable/random habitat classification. In other words, the model predictions are reasonably acceptable as long as they are properly interpreted.

Table 10. Maxent model evaluation summary for k-fold cross validation.

Evaluation statistic	k-fold				
	1	**2**	**3**	**4**	**5**
Presences	45	43	45	45	47
Absences	136	136	136	136	136
AUC	0.64	0.70	0.66	0.61	0.67
Correlation coefficient	0.17	0.28	0.24	0.17	0.27

5 Discussion

5.1 2015 conch abundance survey

The 2015 conch survey results appear to be in line with the trend of recovery and consistently healthy levels of overall Queen conch densities since implemention of a strict management regime in 1994 (Aiken *et al*., 2006; Fisheries Division, 2013). The figure of 409 conch/ha in fact represents the highest mean total conch density recorded since the first survey by Appeldoorn (1995) in 1994, and is also a sizeable increase from the 370 conch/ha reported in 2011 (Fisheries Division, 2013). The stock therefore appears to be well managed in terms of its density and population structure and suggests that the bases on which the integrated management regime of ecosystem-based research and monitoing was well-founded.

In terms of the population structure, overall densities continue to improve but the 2015 survey showed also showed the dynamics of relative density of the various size classes. In this survey for instance the larger size classes, particularly the adults and stoned conch, accounted for most of the total conch abundance and density observed, both overall and within each depth zone (Figures 8 and 9). Total density in 2011 was also found to be driven by juveniles especially by small juveniles (Fisheries Division which prehaps explains the high density of adults observed in this survey as juvenile cohorts' progress into adulthood. Appeldoorn (1995), Tewfik and Appeldoorn (1998) and Smikle and Appeldoorn (2003) reporting on the first three surveys of 1994, 1997 and 2002 noted that total density on the bank was largely driven by juveniles however each by different extents.

The spatial distribution of sites with the highest total density observations favoured the shallower regions to the south and southeast of the bank where the substrate is dominated by seagrass, sand, coral reefs and algal hardground areas (Figure 10). Appeldoorn (1995) and Tewfik and Appeldoorn (1998) also reported that overall densities on the Pedro Bank were highest in algal plains and seagrass habitat in the shallows below 20 m. This is perhaps indicative of the basic substrate complex required by conch whcih are able to adequately facilitate the important activities of feeding and reproduction. In other words this are may represent the intersection of ecological and environmental factors for that time of the year. Regarding the period of the survey which was in October correspond to the period of year (around September) when mature adults conch would be moving from their sandy spawning grounds to habitat associated with their macroalgal food; that is, seagrass, hardsubstrat and coral reef areas (Stoner and Sand, 1992).

High juvenile density on the shallow southeastern edge of the bank was also reported as such in the previous surveys of the Pedro Bank as well (Fisheries Division, 2013). It is perhaps likely that the area is an important nursery site for Queen conch on the bank. Juvenile conch nurseries, as is the case here, are often associated with high density aggregations of juveniles in relatively close proximity to the edges of banks where pelagic larvae preferentially settle then move to their prefered sheltering and feeding habitat (Stoner, 2003; Stoner *et al*., 1996a). The immediate area around the southern edge of the

bank where juvenile conch is highest largely consist of coral reefs (Figure 12) and not seagrass (*Thalassia testudinum*, Banks ex König, 1805) that is often reported as typical nursery habitat in the Caribbean (Stoner and Sandt, 1991). It has been suggested that larval settlement sites and juvenile nursery area may not necessarily be the same as juveniles are believed to make their (largely food-driven) habitat choice after settlement (Stoner, 2003). In the case of the Pedro Bank it would appear that this preference is for sections of seagrass and marcoalgal hardgrounds that are used as nurseries.

Stoner and Sandt (1991) suggested from their study in the Bahamas that algal plains may be unsuitable for juveniles, and in the case of the Pedro Bank would explain the relatively low juvenile abundance and density observed in the deeper depths. Tewfik and Appeldoorn (1998) during the second survey suggested that marcoalgae may have replaced somewhat the use of seagrass by conch as the main nurseries in some of the shallower areas. There may be substance to this as during cross-checking of the Schill habitat map (Baldwin et al., 2014) with field data it was found that there were mant shallow sites (<20 m) that were covered by filatmentous marcoalgae species. *Thalassia testudinum* was not found at any of the 80 sites.

The observed high density juvenile sites are likely such due to site-specific ecological interactions and factors such as tidal retention of larvae which are often difficult to measure. These include factors not necessarily related to just habitat features and adaptive mechanisms aimed at reducing mortality and increasing overall survivorship (Stoner and Lally, 1994).

The observation that both juveniles and mature conch had highest densities in generally similar habitat on the Pedro Bank may be somewhat surprizing (Figures 11 and 12). This given that these shallower habitats are subject to the highest level of conch fishing on the bank. It suggest therefore that some of the negative effects of fishing; namely, the removal of individuals and leaving of dead conch shells on the seafloor may not have had the serious negative effects feared by some. Tewfik and Appeldoorn (1998) cautioned about the possible physiological effects that the presence of dead conch shells could lead to live animals avoiding these areas thus reducing the amount of suitable habitat. In other areas of the region such as in the Bahamas marked reduction in especially adult conch densities have been observed in shallower heavily fished waters less than 10 m (Stoner and Schwarte, 1994). This has not been the case observed on the Pedro Bank where the stock likely benefits from an extremely cautious annual harvest rule of only eight percent of estimated biomass (Smikle, 2010). The stock therefore is able to sustain relatively high densities and support sustained annual fishing pressure.

5.2 Distribution modelling

The importance of selected environmental variables; depth, substrate and the year of the survey, in determining the distribution of mature and juvenile conch on the Pedro Bank was further examined using general additive nixed models (GAMM). The resulting

descriptive realized niche models revealed that depth is indeed important in the overall conch distribution (total conch) particularly driven by mature conch which was statistically significant (Tables 9a-b). The abundance of conch overall (total conch) and mature conch increases with depth according to the model (Figure 23), this as the less vulnerable adults are able to utilize a wider area and also cover a wider spatial range for food and reproductive activity. In the case of juveniles depth is less of a factor perhaps due to their preference for specific nursery areas in the shallow waters (<10 m) (Stoner, 2003).

The models also showed that the nominal variables substrate type and survey year were statistically significant for all groups which suggest that conch abundance on the Pedro Bank changes with respect to the level of each variable. With respect to substrate type seagrss, sand and sediment as well as marcoalgal hardground account for the highest overall abundance (Figure 24). Mature conch abundance was disproportionately higher in seagrass which may be explained by the fact that the survey was conducted at a time of year (October) when adult move in mass from spawning grounds to feeding grounds (Stoner and Sand, 1992). Interestingly though the 2015 survey was conducted in October yet there were relatively large adult abundances in other less productive habitat such as deep coral as well as sand and sediment perhaps still being used for reproductive activity (Brownell and Stevely, 1981). Scientific divers conducting the 2015 survey reported observations of reproductive behavior; that is, coupling, mating and egg masses on the substrate and in the water column. The most important substrates for juvenile abundance according to the model were seagrass, sand and sediment, deep coral closely followed marcoalgal hardground areas all of which are normally associated with ample amounts of food.

5.3 Predictive models

Predictive habitat maps of the potential abundance of Queen conch on the Pedro Bank were also produced from the results of the GAMMs. The models generally revealed that there is a relatively high potential for high overall abundance of the species across much the bank particularly areas below the 20 m depth contour (Figure 27). These areas correspond largely with high density areas from the results of the 2015 which lay to the east and southeastern regions of the dominated by seagrass, sand and sediment, and macroalgal hardground substrate. The areas of lowest predicted abundance according to the model were those to the northwest and around most of the periphery of the bank.

The areas of highest potential abundance for mature conch were shown to be very distinct areas that corresponded to areas of seagrass (mixed with macroalgae) surrounded by sand and sediment areas located to the southeast of the bank (Figure 28). It should be noted from the previous section however that mature conch also had fairly high aggregations in other areas and on other substrates as well. It is perhaps the substrate complexity as opposed to any one substrate type, which is favoured by mature conch and brought out in the model predictions. In other words, it is the proximity of feeding grounds (marcoalgae and seagrass) and spawning grounds (sand and sediment) that is the important driving factor for mature conch distribution in the shallow areas of the Pedro Bank.

The predictive habitat map for juveniles showed that potential abundance was highest in areas of sand and sediment which covers most of the substrate below the 20 m depth

contour as well as large sections of the extreme west and northwest of the bank (Figure 29). The potential for juvenile nurseries occurring in these deeper and remote parts of the bank is somewhat surprising and encouraging as well. These areas may represent future management areas that can facilitate protected area designation for conch or at the very least lead to further exploration of their use in managing the stock. Another very important revelation of the map is that the spatial extent of potential high abundance (potential conch nurseries) is much larger than the observed juvenile distribution. Only a relatively small portion of potentially "optimal" habitat was shown to have large aggregations of juveniles (Figure 12). This further serve to underscore the complexity of distributional factors at work in juveniles and the need to properly manage and protect these preferred nursery areas. Stoner *et al*., (1996b) in their study of juvenile distribution in the Bahamas also found that large aggregations of juveniles occupied only a small proportion (1.5%) of potentially suitable habitat.

An additional prediction was made of the total conch distribution using the Maxent programme in order to estimate habitat suitability of the bank. The Maxent was used here: (i) to have another prediction algorithm for comparison with the GAMM predictions, (ii) to incorporate an additional variable (chlorophyll-a), (iii) estimate the extent of potentially suitable habitats, and (vi) explore the effectiveness of Maxent as a management tool given the high cost associated with obtaining abundance data. Featuring as many covariates as possible was important as simple variables such as substrate and depth alone may not be sufficient to explain conch distribution (Stoner, 2013). The model included the variables depth, substrate and chlorophyll-a concentration of the water column which was included as a proxy variable for primary productivity, an important factor in the availability of the species' macroalgal food. The resulting habitat suitability predictions favour similar areas indentified as high probabily of abundance by the GAMM predictions (Figure 30). The least suitable habitats were found to occur in much of the deeper northwestern areas of the bank. There were a few differences however as as the maxent predictions indicated that there were areas of high potential suitability in a few isolated areas to the extreme east and on the extreme southern tip of the bank. This difference is brought out as a result of the additional variable, chlorophyll-a, which the model ranked as more important to its' perdiction than the other covariates; depth and substrate type.

6 Conclusions and management implications

The relationship between habitat and species abundance is fundamental to species distribution models (SDM) however this relationship can rarely be tested with absolute confidence as the exact requirements of a species is usually poorly understood (Stoner, 2003), and attempting to measure or even quantify these requirement are either impractical or not cost-effective (Robinson *et al.*, 2011). Despite these limitations, the use of SDMs here has shown yet again its usefulness in facilitating a better understanding of species-environment relationships as well as identifying important knowledge gaps that may be cruial in the the management of Queen conch on the Pedro Bank. These are important given the numerous threats faced by species' partticularly habitat loss and the direct removal of individuals through fishing. Futunately there is a concerted drive by the authorities in Jamaica to address these threats through sound ecosystem based management (ESBM). The findings here are relevant to this management approach and may be fed directly into the general fisheries management of the species as well as other initiatives for the conservation of biodiverity on the Pedro Bank.

From the results of the models it is clear that the factors determining the distribution of Queen conch over space and time on the Pedro Bank are very complex. These factors likely represent an intercept among various ecological, environmental and biological processes that are site-specific. There were many patterns observed here that did not those observed in other areas of the Caribbean and also instances where the species did not occupy seemingly optimal environments. This is probably due to unquie differences in habitat/substrate structure of the bank as well as local differences in food availability and predation for example.

In terms of management therefore the focus should be on site-specific and conch-specific approaches that consider environmental variables used here as well as other potentially important factors affecting the species and the quality of its habitat. This is opposed to general conservation strategies or strategies developed from different localities. This approach will hopefully increase tremendously the explanatory power of the models and provide an even greater understanding of the species-environment relationship that can guide specific management decision and action. Other factors also be incorporated if possible including, but not limited to: (i) direct and reliable measures of fishing effort and other anthropogenic causes of habitat change, (ii) more precise temperature measurements, (iii) measures of the relative size of different substrate utilized by the population, and (iv) measures of larval settlement and recruitment. Ideally, these variables should also incorporate a dynamic or temporal aspect which goes further than just collecting static data such as from surveys or remote sensing. The management approach should additionally incorporate important biological and ecological processes if possible such as the daily and seasonal migrations of the species between different habitat types for the purposes of feeding and reproduction.

It is however recognised that there are many issues involved in the implementation of such an intensive management strategy considering Jamaica is a small island developing state having to deal with prolonged resource limitations and other major competing priorities. Regardless, this a worthwhile strategic direction for the species' sustainable management and development on the Pedro Bank and more general management plans considering its various other resources and stakeholders.

In terms of priority areas for improving the ecosystem based management (ESBM) regime for the species, it would be a sensible approach to focus on conservation and preservation of those high density areas as revealed by the descriptive models as well as those high potentially suitable and abundant habitats as revealed by predictive models. Information from these areas should help to form the basis of future management of the species and the relationship with its' habitat. Some suggested specific uses of this information from this study includes:

(i)   Incorporation into the on-going Pedro Bank marine spatial planning (MSP) process and similar plans aimed at the establishment of protected areas and management of Jamaica's coastal and marine zones.

(ii)   Guiding the design/improvement of conch-specific and site-specific habitat monitoring programme aimed at conserving critical nursery areas, spawning grounds and feeding areas.

(iii)   Incorporation into a revised and updated conch fishery management plan for the Pedro Bank.

(iv)   Guide the design/improvement and prioritization of fishery management areas for conch aimed at, among other things, ensuring minimal negative impact of fishing activity on their habitat and also ensuring that high-density aggregations are not depleted.

(v)   Add to the national marine spatial literature databases for use by public and private bodies in the furtherance of sustainable ESB management of Jamaica´s marine resources.

.

References

Aiken, K, Kong, A, Smikle, S, Appledoorn, R., and Warner, G. (2006). Managing Jamaica's Queen Conch Resources. Ocean and Coastal Management, 49, 332–41.

Aiken, K. A., Kong, G. A., Smikle, S., Mahon, R., and Appeldoorn, R. (1999). The queen conch fishery on Pedro Bank, Jamaica: discovery, development, management. Ocean and Coastal Management, 42(12), 1069–1081.

Akaike, H. (1973). Information theory as an extension of the maximum likelihood principle. In: Petrov, B.N., Csaki, F. (Eds.), Second International Symposium on Information Theory. Akademiai Kiado, Budapest, Hungary, 267–281.

Allen, R. and Webber, D. (2013). The fishers of the Pedro Bank, Jamaica: through the lens of their livelihoods. Global Journals of Human Social Science, 13(7), 43-50.

Appeldoorn, R. S. (1988). Age determination, growth, mortality and age of first reproduction in adult queen conch, *Strombus, gigas* L., off Puerto Rico. Fisheries Research 6, 363–378.

Appeldoorn, R.S. 1995. Stock abundance and potential yield of the queen conch of Pedro Bank. Univ. Puerto Rico/Fisheries Div. Jamaica. 30.

Baldwin, K. and R. Mahon. (2015). Rapid Fisheries Sector Assessment of the Jamaican Seascape. For The Nature Conservancy, Caribbean Marine Biodiversity Project. USAID, Kingston, Jamaica. 106.

Baldwin, K., Schill, S., Zenny, N., and Blake, D. (2014). Developing ecosystem-based information for marine spatial planning on the Pedro Bank, Jamaica. 67th Gulf and Caribbean Fisheries Institute, Barbados.

Becker, R. A., Allan R. Wilks, A. R., Ray Brownrigg, R. and Minka, T. P. (2014). maps: Draw Geographical Maps. R package version 2.3–9. http://CRAN.R-project.org/package=maps.

Bivand, R. S., Pebesma, E. and Gomez-Rubio, V. (2013). Applied spatial data analysis with R, Second edition. Springer, NY.

Brownell, W. N. and J. M. Stevely. (1981). The biology, fisheries, and management of the queen conch, *Strombus gigas*. Marine Fisheries Review, 43(7), 1–12.

Brunsdon, C. and Chen, H. (2014). GISTools: Some further GIS capabilities for R. R package version 0.7-4. http://CRAN.R-project.org/package=GISTools.

Board, M. A. (2005). Millennium ecosystem assessment. Washington, DC: New Island.

Borcard, D., Gillet, F., and Legendre, P. (2011). Numerical ecology with R. Springer Science and Business Media.

Canty, A. and Ripley, B. (2014). boot: Bootstrap R (S-Plus) Functions. R package version 1.3–13.

Carpenter, G., Gillison, A. N., Winter, J. (1993). DOMAIN: a flexible modelling procedure for mapping potential distributions of plants and animals. Biodiversity Conservation, 2, 667–680 pp.

Chakallal, B. and K. L. Cochrane. (1996). The queen conch fishery in the Caribbean – An approach to responsible management. Pres. At Int'l. Mtg. On Queen Fishery, Puerto Rico, July 29–31, 18.

CRFM. (2012). Report of Eighth Annual Scientific Meeting – Kingstown, St. Vincent and the Grenadines, 20 – 30 June 2012. CRFM Fishery Report – 2012, 1, 27 pp.

Davison, A. C. and Hinkley, D. V. (1997). Bootstrap Methods and Their Applications. Cambridge University Press, Cambridge. ISBN 0-521-57391-2.

Dolan, P. 1972. Genesis and distribution of recent sediments on Pedro Bank, south of Jamaica. Ph.D thesis, University College, London. 232 pp.

Elith, J., Graham, C.H., Anderson, R.P., Dudík, M., Ferrier, S., Guisan, A., Hijmans, R. J., Huettmann, F., Leathwick, J. R., Lehmann, A., Li, J., Lohmann, L. G., Loiselle, B. A., Manion, G., Moritz, C., Nakamura, M., Nakazawa, Y., Mc Overton, J. C. M., Townsend, P. A., Phillips, S. J., Richardson, K., Scachetti-Pereira, R., Schapire, R. E., Soberón, J., Williams, S., Wisz, M. S. and Zimmermann, N. E. (2006). Novel methods improve prediction of species' distributions from occurrence data. Ecography 29, 129–51.

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

Elith, J., Phillips, S. J., Hastie, T., Dudík, M., Chee, Y. E., and Yates, C. J. (2011). A statistical explanation of MaxEnt for ecologists. Diversity and Distributions, 17(1), 43–57.

Efron, B and Tibshirani, R.J. (1993). An Introduction to the Bootstrap. Chapman and Hall/CRC.

Ehrhardt, N. M., and Valle-Esquivel, M. (2008). Conch (*Strombus gigas*) stock assessment manual. Caribbean Fishery Management Council.

Fisheries Division. (2013). Pedro Bank 2011 conch Abundance Survey Report [Unpublished/manuscript]. Fisheries Division, Ministry of Agriculture and Fisheries, Jamaica.

Fisheries Division. (2014). Fisheries Division, Ministry of Agriculture and Fisheries, Jamaica.

Fitzmaurice, G. M., Laird, N. M. and Ware, J. H. (2004). Applied Longitudinal Analysis. John Wiley and Sons, Hoboken, NJ.

GES-DISC/NASA. (2014). Ocean Color Online Visualization and Analysis System using Giovanni. [URL: http://disc.sci.gsfc.nasa.gov/giovanni/]

Guisan, A., Edwards, T. C., and Hastie, T. (2002). Generalized linear and generalized additive models in studies of species distributions: setting the scene. Ecological Modelling, 157(2), 89‑100.

Guisan, A., and Zimmermann, N. E. (2000). Predictive habitat distribution models in ecology. Ecological Modelling, 135(2), 147–186.

Hastie, T. J., Tibshirani, R. J. (1990). Generalized Additive Models. Chapman & Hall.

Hijmans R. J. (2015). raster: Geographic Data Analysis and Modeling. R package version 2.3–40. http://CRAN.R-project.org/package=raster.

Hijmans, R. J., Phillips, S., Leathwick, J. and Elith, J. (2015). dismo: Species Distribution Modeling. R package version 1.0–12. http://CRAN.R-project.org/package=dismo.

Huang, Z., Brooke, B., and Li, J. (2011). Performance of predictive models in marine benthic environments based on predictions of sponge distribution on the Australian continental shelf. Ecological Informatics, 6(3), 205–216.

Intergovernmental Oceanographic Commission (IOC), International Hydrographic Organisation (IHO), and British Oceanographic Data Centre (BODC). (2003). Centenary Edition of the GEBCO Digital Atlas [CD-ROM], Br. Oceanogr. Data Centre, Liverpool, U. K.

Kaiser, M. J., Collie, J. S., Hall, S. J., Jennings, S., and Poiner, I. R. (2003). 12 Impacts of Fishing Gear on Marine Benthic Habitats. Responsible fisheries in the marine ecosystem, 57(3), 197.

Magris, Rafael A., and Déstro, G. F. G. (2010). Predictive modeling of suitable habitats for threatened marine invertebrates and implications for conservation assessment in Brazil. Brazilian Journal of Oceanography, 58(spe4), 57–68. Retrieved December 09, 2015, from http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1679-87592010000800008&lng=en&tlng=en.

Malanson, G.P., Westman, W.E., Yan, Y. L., 1992. Realized versus fundamental niche functions in a model of chaparral response to climatic change. Ecological Modelling, 64, 261–277.

Marshak, A. R., Appeldoorn, R. A, and Jimenez, N. 2006. Utilization of GIS mapping in the measurement of the spatial distribution of Queen conch (*Strombus gigas*) in Puerto Rico. GCFI: 57.

Merow, C., Smith, M. J., and Silander, J. A. (2013). A practical guide to MaxEnt for modeling species' distributions: what it does, and why inputs and settings matter. Ecography, 36(10), 1058–1069.

Moore, C. H., Harvey, E. S., and Van Niel, K. (2010). The application of predicted habitat models to investigate the spatial ecology of demersal fish assemblages. Marine Biology, 157(12), 2717-2729.

Munro, J. L. and Thompson, R. (1983). Ch 3: Areas Investigated, Objectives, and Methodology. In: Munro, J. L. (Ed.), Caribbean Coral Reef Fishery Resources. ICLARM Studies and Reviews 7, 15–19.

Nix, H.A. (1986). A Biogeographic Analysis of Australian Elapid Snakes. In: Longmore, R. (Ed.), Atlas of Elapid Snakes of Australia. Australian Flora and Fauna Series Number 7. Australian Government Publishing Service, Canberra.

Pante, E. and Simon-Bouhet, B. (2013). marmap: A Package for Importing, Plotting and Analyzing Bathymetric and Topographic Data in R. PLoS ONE 8(9): e73051. doi:10.1371/journal.pone.0073051

Pebesma, E. J. (2004). Multivariable geostatistics in S: the gstat package. Computers and Geosciences, 30, 683–691.

Phillips, S.J., Anderson, R.P., and Schapire, R.E. (2006). Maximum entropy modeling of species geographic distributions. Ecological Modelling, 190, 231–59.

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

Pinheiro, J.C., and Bates, D.M. (2000). Mixed-Effects Models in S and S-PLUS. Springer.

Pinheiro, J. C, Bates, D.M, DebRoy, S., Sarkar, D. and R Core Team (2014). nlme: Linear and Nonlinear Mixed Effects Models. R package version 3.1–118, <URL: http://CRAN.R-project.org/package=nlme>.

Prada, M., Castro, E., Taylor, E., Puentes, V., Appeldoorn, R., and Daves, N. (2008). Non-detrimental findings for the Queen Conch (*Strombus gigas*) in Colombia. In *NDF Workshop Case Studies WG*. [Retrieved from https://cites.unia.es/file.php/1/files/WG9-CS3.pdf.]

QGIS Development Team. (2015). QGIS Geographic Information System. Open Source Geospatial Foundation Project. [http://qgis.osgeo.org]

R Core Team. (2014). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Raven, P. H. (2002). Predicting species occurrences: issues of accuracy and scale. J. M. Scott, P. Heglund, and M. L. Morrison (eds.). Island Press.

Reiss, H., Birchenough, S., Borja, A., Buhl-Mortensen, L., Craeymeersch, J., Dannheim, J., and Degraer, S. (2014). Benthos distribution modelling and its relevance for marine ecosystem management. ICES Journal of Marine Science, 72, 297–315.

Robinson, L. M., Elith, J., Hobday, A. J., Pearson, R. G., Kendall, B. E., Possingham, H. P. and Richardson, A. J. (2011). Pushing the limits in marine species distribution modelling: lessons from the land present challenges and opportunities. Global Ecology and Biogeography, 20, 789–802.

Sarkar, D. (2008). Lattice: Multivariate Data Visualization with R. Springer, New York. ISBN 978-0-387-75968-5.

Schabenberger, O. and Pierce, F. J. (2002). Contemporary statistical models for the plant and soil sciences. Boca Raton, FL: CRC Press.

Smikle, S. G. (2010). Pedro Bank Queen Conch fishery, assessment and TAC recommendation. Unpublished report of the Fisheries Division of Jamaica, Ministry of Agriculture.

Smikle, S. G and Appeldoorn, R. S. (2003). 2002 Estimates of Abundance and Potential Yield for the Pedro Bank Queen Conch Population. Fisheries Division Ministry of Agriculture Kingston, Jamaica.

Stoner, A. W. (2003). What constitutes essential nursery habitat for a marine species? A case study of habitat form and function for queen conch. Marine Ecology Progress Series, 257, 275–289.

Stoner, A. W., Glazer, R. A., and Barile, P. J. (1996a). Larval supply to queen conch nurseries: relationships with recruitment Process and population size in Florida and the Bahamas. Journal of Shellfish Research, 15(2), 407–420.

Stoner, A. W., Hanisak, M. D., Smith, N. P., and Armstrong, R. A. (1994). Large-scale distribution of queen conch nursery habitats: implications for stock enhancement. In: Appledoorn RS, Rodriguez B (eds) Queen conch biology, fisheries and mariculture. Fundacion Científica Los Roques, Caracas, 169–189.

Stoner, A. W., Pitts, P. A., and Armstrong, R. A. (1996b). Interaction of physical and biological factors in the large-scale distribution of juvenile queen conch in seagrass meadows. Bulletin of Marine Science, 58(1), 217–233.

Stoner A. W. and Lally. J. (1994). High-density aggregation in queen conch Strombus gigas: formation, patterns, and ecological significance. Marine Ecology Progress Series, 106, 73–84.

Stoner, A. W. and Sandt, V. J. (1991). Experimental analysis of habitat quality for juvenile queen conch in seagrass meadows. Fishery Bulletin, U.S. 89, 693–700.

Stoner, A.W. and Sandt, V. J. (1992). Population structure, seasonal movements and feeding of queen conch, *Stombus gigas*, in deep-water habitats of the Bahamas. Bulletin of Marine Science 51, 287–300.

Stoner, A. W., and Ray-Culp, M. (2000). Evidence for Allee effects in an over-harvested marine gastropod: density-dependent mating and egg production. Marine Ecology Progress Series, 202, 297–302.

Stoner, A. W., and Ray, M. (1996). Queen conch, Strombus gigas, in fished and unfished locations of the Bahamas: Effects of a marine fishery reserve on adults, juveniles, and larval production. Fishery Bulletin, 94, 551–556.

Stoner, A. W. and Schwarte, K. C. (1994). Queen conch, *Strombus gigas*, reproductive stocks in the central Bahamas: distribution and probable sources. Fishery Bulletin, 92,171–179.

Tewfik, A. (1996). An assessment of the biological characteristics, abundance, and potential yield of the queen conch (*Strombus gigas* L.) fishery on the Pedro bank off Jamaica. M.Sc. thesis, Acadia University, Wolfville, NS, 139.

Tewfik, A. and R.S. Appeldoorn. (1998). 1997 Queen Conch (*Strombus gigas*) Abundance Survey and Potential Yield Estimates for Pedro Bank, Jamaica. Fisheries Division, Jamaica.

Wintle, B. A, and Bardos, D. C. 2006. Modelling species habitat relationships with spatially autocorrelated observation data. Ecological Applications 16, 1945–58.

Wood, S. N. (2004) Stable and efficient multiple smoothing parameter estimation for generalized additive models. Journal of the American Statistical Association, 99, 673–686.

Wood, S. N. (2006) Generalized Additive Models: An Introduction with R. Chapman and Hall/CRC.

Yoder, J. (2013, April 23). Species distribution models in R (The Molecular Ecologist). [Retrieved January 4, 2016 from http://www.molecularecologist.com/2013/04/species-distribution-models-in-r/]

Zuur, A., Ieno, E. N., Walker, N., Saveliev, A. A., and Smith, G. M. (2009). Mixed effects models and extensions in ecology with R. Springer Science and Business Media.

Zuur, A. F., Ieno, E. N. and Elphick, C. S. (2010). A protocol for data exploration to avoid common statistical problems. Methods in Ecology and Evolution, 1, 3–14.