MULTIVARIATE HABITAT-BASED PREDICTIVE MODELING OF THREE DEMERSAL ROCKFISH SPECIES IN CENTRAL CALIFORNIA

By

Georgia Rae Martel

A Thesis Presented to
The Faculty of Humboldt State University
In Partial Fulfillment of the Requirements for the Degree
Master of Science in Biology

Committee Membership
Dr. Brian Tissot, Committee Chair
Dr. Andre Buchheister, Committee Member
Dr. James Lindholm, Committee Member
Dr. Timothy Mulligan, Committee Member
Dr. Paul Bourdeau, Program Graduate Coordinator

May 2020

ABSTRACT

MULTIVARIATE HABITAT-BASED PREDICTIVE MODELING OF THREE DEMERSAL ROCKFISH SPECIES IN CENTRAL CALIFORNIA

Georgia Rae Martel

Accurate, spatially explicit models of rockfish abundance are critical in implementing ecosystem-based management strategies and designating essential fish habitats and marine protected areas. Multibeam bathymetry methods and visual, non-extractive submersible transect surveys were combined to collect environmental variables and fish abundance data at three distinct sites within the study region. Zero-adjusted models were developed using habitat classification analyses of high-resolution (5 m) digital elevation models. Model accuracies were assessed by using a reserved subset of the original datasets. To demonstrate that a model's predictive power was linked to its spatial origins, Mean Absolute Error and coefficient of determination values were recorded when the site-trained model was used to predict that site's own test data, and when it predicted species' distribution at the two other sites whose training data were not used to inform the model. The habitat characteristics of importance to each species varied across sites, and model accuracies declined when applied to a site that differed in physical composition, suggesting a species will alter their habitat associations in accordance to the relative availability of preferred substrata and terrain.

ACKNOWLEDGEMENTS

I thank the Council on Ocean Affairs, Science, and Technology's Graduate Student Research Grant Award for funding this research. My advisor Dr. Brian Tissot for his continued guidance and patience. Dr. Timothy Mulligan for his support and for agreeing to be on my committee despite enjoying retirement. Dr. Andre Buchheister for his invaluable assistance as a statistical advisor. Dr. Mark Henderson for his code and instruction on accounting for spatial autocorrelation. Dr. James Lindholm for his advice and being on my committee. The entire *Delta* team that collected these data, including Dr. Rick Starr and Mary Yoklavich. Dr. Diana Watters for compiling the data I requested and use of her DEM mosaic. I would also like to thank my lab cohort, Jana Litt, Carolyn Belak, Christopher Teague, and Adam Copeland, for their amazing support and camaraderie. And I thank my husband, Chad Martel, a better partner than I ever could have asked for, for keeping me sane through it all.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

INTRODUCTION

Many populations of commercially and recreationally important groundfish species, including rockfish in the genus *Sebastes*, have long been in decline along the Pacific northwest of the United States, prompting concern from managers and communities whose livelihoods depend on these stocks (Starr et al. 2002; Harvey et al. 2006; Levin et al. 2006). Rockfish are one of the most economically valuable commercial fisheries on the west coast, in addition to being important recreational fisheries (Yoklavich et al. 2000; OCNMS 2011). This taxon presents a challenge to traditional management strategies however because of their unique life history traits. Rockfishes are slow growing, long-lived, and reach age of maturity much later in life than many other commercial stocks, rendering these populations especially vulnerable to overfishing (Love et al. 1990; Parker et al. 2000; Yoklavich et al. 2000; Levin et al 2006).

Recently, managers have been seeking forms of human impact mediation and ecosystem-based management strategies to combat the declines seen in many rockfish populations (PFMC 2010). Investigating the factors that affect the degree to which fishes use specific habitat characteristics will aid managers' decisions regarding allocation of resources. The revisions to the Magnuson-Stevens Fishery Conservation and Management Act (MSFCMA) allowed for the inclusion of habitat in fishery management strategies (MSA, 16 U.S.C. § 1801 et seq.). The MSFCMA revisions provided an option for decision-makers to designate certain areas as Essential Fish Habitat (EFH) and Habitat Areas of Particular Concern (HAPC) (Morgan et al. 2005; NOAA 2010). EFH is defined as “those waters and substrate necessary to fish for spawning, breeding, feeding or growth to maturity” (Rosenburg et al. 2000). HAPCs are areas within EFH that are rare and are either ecologically important, sensitive to disturbance, or may be stressed (EA 2013). While some aspects of EFH have been defined and protected for federally managed fishes, many rockfish species’ habitat associations are still uncertain, and the question of whether there are discernable ontogenetic patterns in species’ habitat use remains. Regarding the species for which EFH has been defined, it remains unclear whether these associations hold constant throughout these species’ ranges. Disagreement also exists around what constitutes “essential” habitat for these fishes, with multiple studies finding evidence of facultative habitat usage by rockfish species (Auster et al. 1995; Auster 2005; Wrubel 2013).

In order for marine spatial planning strategies such as the designation of EFH, HAPC and the implementation of Marine Protected Areas (MPAs) to be successful, a robust understanding of the habitat characteristics of importance to the species of interest is required. Additionally, there is a need to understand how these associations with habitat may vary across broad geographic scales. Research in this field is sparse, with many studies focusing only on limited geographic regions and often reporting seemingly differing findings (Barrs et al. 1992; Auster et al. 1995; Yoklavich et al. 2000; Pacunski et al. 2001; Johnson et al. 2003; Anderson and Yoklavich 2007; Love et al. 2009). To strengthen inferences about any specific rockfish stock, multiple sub-populations across a number of sites that are diverse in physical composition should be compared. However, rockfish’s affinity for high relief, rocky reefs makes ascertaining habitat data difficult using conventional methods such as bottom trawls (Wedding and Yoklavich 2015). Non-extractive visual methods such as submersible and Remotely Operated Vehicle (ROV) surveys are a viable and effective means of accurately sampling the benthic environment (Auster and Lindholm 2005; Yoklavich et al. 2007; Tissot et al. 2008; Anderson et al. 2009; Knight et al. 2014; Lindholm et al. 2015). When combined with the recent advancements in remote sensing technologies and predictive modeling approaches, visual sampling methods provide a powerful tool for researchers to effectively predict the distribution of rockfish species, without the need for exhaustive sampling (Summers-Morris et al. 2004; Iampietro and Kvitek 2005; Young 2007; Iampietro et al. 2008; Young et al. 2010; Ierodiaconou et al. 2011). The information gathered on species' abundance in relation to a suite of environmental variables (slope, aspect, elevation, etc.) can be analyzed with statistical models such as Generalized Linear Models (GLMs) to predict where species are most likely to occur in areas where no direct observational data exists (Rotenberry et al. 2006).

In this study, I investigated the limitations and potential of a relatively new class of habitat-based predictive distribution models when applied across a large swath of coastline. My goals were (1) to identify the habitat characteristics associated with observations of three demersal rockfish species; (2) test whether these habitat associations remained constant at three different sites that varied in habitat structure and availability; and (3) evaluate these differences in habitat usage using predictive modeling techniques. To compare habitat associations between sites, I developed species-specific models that predicted the distribution of each rockfish species at each of my three study sites, and then I assessed the accuracy of these site-trained models when predicting the same species' distribution at other sites that differed in physical composition, with the expectation being that model fit would decrease when applied to areas with habitat features dissimilar to the original site of model construction.

To allow for comparison, I selected species of rockfish that were common to all three survey sites in my study: Squarespot (*Sebastes hopkinsi*), Rosy (*S. rosaceus*), and Pygmy (*S. wilsoni*) rockfish (Eschmeyer 1983). Based on the current literature, I had reason to expect these species' habitat associations to be similar but with nuanced differences (Love et al. 2002). Rosy rockfish have been shown to prefer complex substrata, positioning themselves in the upper depths of rocky reefs, between small cracks and crevices. Squarespot rockfish associate strongly with the same type of substrate, but at mid-depths on the bank or reef (Anderson et al. 2009). While squarespots are found in greater abundances over high relief structure, they have also been observed associating with lower complexity substrata such as cobble (Laidig et al. 2009), revealing a more generalist habitat selectivity behavior. As a schooling species, Pygmy rockfish have typically been found to associate with shallower waters relative to Rosy and Squarespot habitats, and can be found over a multitude of substrata, including sand and mud (Anderson et al. 2009). The nuanced differences in each species' microhabitat associations amplified my ability to detect geographic changes in habitat usage among species. The data used in this study were gathered from baseline surveys in the Monterey Bay National Marine Sanctuary (MBNMS) conducted in 2007 directly following the designation of State MPAs under the California Marine Life Protection Act (MPLA) (Starr et al. 2008).

Analyzing these baseline data presents an opportunity to examine fish-habitat associations before any MPA protections were in effect and provides a reference point to compare any changes to these relationships in the future. Because these species are not specifically targeted in commercial or recreational fisheries, they already benefit from some protection, although they are subjected to varying levels of bycatch. Due to this de facto protection, these species have rarely been studied in the context of MPA management. Therefore, we do not understand the effects MPA closures will have on these species. The results of this project provide contextual information on how distinct species of fish utilize habitat at varying locations within their geographic range and shed light on the limitations of applying predictive distributional models to large swaths of coastline. Better understanding of these complex ecological relationships will provide policy-makers with the ability to make better-informed decisions and improve the management of our ocean and its resources.

METHODS

Study Region

The data used in this project were collected during the 2007 submersible *Delta* surveys off the coast of central California, at the time of the initial establishment of MPAs. These surveys encompassed 164,000 m^2 of seafloor ranging in depth from 24 to 365 m (Figure 1). Surveys took place largely within state waters (three nautical miles from shore), and encompassed a broad region of the continental shelf within the Monterey Bay marine sanctuary (Starr et al. 2008). The waters in this region are productive and cool, owing to the California Current's year-round equatorward direction of flow, and substantial spring and summer upwelling events (Hickey 1998), which supports diverse fish and invertebrate communities. The area also historically sustained large commercial and recreational fisheries (Miller and Geibel 1973; Karpov et al. 1995; Mason 1998).

Data Collection

To compare habitat associations of targeted fish species, a total of 20 dives and 45 transects from three distinct study sites, Portuguese Ledge, Pt. Lobos, and Pt. Sur, within the survey region were selected to encompass the known occurrence depths of the three study species (Figure 2). The purpose of selecting these three sites was to account for the strong influence of depth on structuring fish assemblages, while comparing fish habitat usage in areas with varying substrate profiles.

These sites exhibit a wide range of benthic habitats, from sloping mudflats to high relief rocky pinnacles and boulder fields (Wedding and Yoklavich 2015). Portuguese Ledge is a deep mudplain just south of the Monterey Canyon. This region is characterized by gently sloping mudflats and a low availability of high relief rocky reef. This was the deepest site surveyed, averaging about 100m in depth across all transects. Point Lobos, in contrast to Portuguese Ledge, has abundant high relief structure and high rock pinnacles. Pt. Lobos is extremely productive and supports a high diversity of invertebrate populations. This was the shallowest site sampled, averaging about 45m. The southern-most site, Point Sur, can be characterized mainly by boulder fields with some high relief reefs dispersed throughout. Pt. Sur averaged about 70m in depth across transects.

As part of the *Delta* protocol, quantitative, visual strip transects were conducted to characterize seafloor habitat and identify, count, and measure species of fishes. Each dive included two to four 10 min-long transects along a predetermined habitat-depth strata based on the occurrence of rocky habitat identified from multibeam bathymetric maps provided by the California Seafloor Mapping Program (SFML 2006). A pilot operated the submersible while an experienced scientist identified and estimated total length of all fish species using paired lasers spaced 20 cm apart as a guide within a 2 m-wide strip adjacent to the submersible. The length of each dive was determined using a Doppler velocity log (DVL) (NavQuest 600 Micro) and a ring-laser gyrocompass (Starr et al. 2008). The width of each transect was estimated at 2 m by keeping the submersible at a constant altitude of 1-2 m above the benthos (Yoklavich et al. 2002).

Each transect was documented with an external video camera and annotated in real-time by the scientific observer on board. These baseline surveys were conducted in September-November of 2007, the same year these areas were given State protection. For this reason, protections had little effect on the stock at this time, therefore MPA boundaries such as the delineations between State Marine Reserves and State Marine Conservation Areas were not considered in my analyses.

Video Analysis

From the annotated video, the geographic coordinates of the target fishes on transects were recorded when the individual fish was positioned between the submersible’s paired lasers in the center of the viewing area. The purpose of collecting these detection data this way was to standardize fish observations between the three species, which exhibit different behavioral responses to external stimuli such as the presence of a large submersible (Stoner et al. 2009). These georeferenced presence/absence data were converted into the ESRI shapefile format for GIS analysis. If the fish did not move within the paired lasers, its georeferenced timestamp was recorded when it was nearest the center of the viewing area. If more than one individual of a species was observed within a single second (a moving school for example) then the video was paused when more than half of the individuals were nearest the center of the viewing field. The total number of individuals in the group was estimated and their location was recorded as a single set of coordinates. Because these three species of rockfish are known to exhibit strong site fidelity, sampled transects were considered to be independent.

Fish sizes were estimated and binned at 5 cm intervals (Wedding and Yoklavich 2015). Total length at age of maturity (Haldorson and Love 1991) was utilized to delineate three age classes of each of the target species to determine whether there were any ontogenetic patterns of habitat association; fish <10 cm were classified as young-of-the-year (YOY), fish in the 10-15 cm bin were classified as sexually immature adolescents deemed “Year-1s”, and fish ≥ 15 cm were classified as reproductive adults (“Year-2s”). Each species’ age class data were pooled by site for initial statistical analyses and model-building. Species’ age classes were later modeled separately to determine any influence of ontogenetic patterns of habitat association if the initial pooled models did not converge.

Substratum (mud, sand, gravel, pebble, cobble, boulder, flat rock, rock ridge, and pinnacle top) was classified from the recorded video of each transect based on standard geological definitions (Stein et al. 1992; Greene et al. 1999). Distinct habitat patches were delineated as such if the substrate type remained constant for at least three seconds during video playback (Wedding and Yoklavich 2015). A patch was comprised of a primary substrate (defined as occupying at least 50% of the area viewed), and a secondary substrate (>20% of the area viewed). The length of each habitat patch was determined using the starting and end coordinates of each distinct patch.

Relief was categorized based on the complexity of the primary and secondary substrata. I simplified this categorical variable into four levels using similar terminology as Wedding and Yoklavich (2015). The highest relief category was large structured hard substratum (Lhard), which included pinnacle tops, rock ridges, and flat rocks. Boulders, cobble, and pebble were classified as moderate complexity hard substratum (Mhard), while sand and mud comprised the soft substratum (Soft) low complexity category. If the primary substrate was of a different relief category than the secondary substrate (boulders scattered across a mud plain for instance), then the resulting relief was coded as “Mix”.

Geospatial Processing

High-resolution multibeam data were collected across the entire study region by the California Seafloor Mapping Program at a 5 m resolution. A hillshade layer was created and overlaid onto the Digital Elevation Model (DEM) so that physical features such as rocky reefs and outcroppings were more apparent. After confirming the spatial reference system (World Geodetic System 1984), the submersible transect data were projected onto the bathymetric DEM as points in ArcMap (Esri Inc. 10.5.1) whereby each point corresponded to the geographic coordinates of the submersible recorded by the DVL at one-second intervals.

Following preparation of the coordinate data, slope and aspect were derived using the ArcGIS Spatial Analyst extension. Previous studies have shown that rockfish prefer sloping terrain (McClatchie et al. 1997). Consequently aspect, measured as the compass direction a slope faces, was hypothesized to be an important predictor of rockfish abundance due to the relationship between cardinal direction of slope face and the delivery of nutrients from the prevailing current (Young et al. 2010). The three other habitat covariates were derived from the DEM using the Benthic Terrain Modeler extension (Walbridge et al. 2018). Bathymetric Positioning Index (BPI) is a habitat metric that characterizes a pixel in a bathymetric DEM as a positive or negative feature relative to the surrounding seafloor elevation (Weiss 2001; Young et al. 2010). Locations that are higher than their surroundings will have positive values, whilst areas that are lower will have negative values. Flat areas have values close to zero (Ierodiaconou 2011). BPI is a useful tool in examining habitat associations, and it can be calculated at different spatial extents. In this study, I used a fine-scale BPI (f-BPI), calculated at a scale factor of 36 m, and a broad-scale BPI (b-BPI) at a scale factor of 200 m. Since these species of rockfish exhibit generally high site fidelity, I expected f-BPI to be a better predictor of occurrence as this scale captures the finer microhabitats individual fish are most likely associating with. BPI at both scales was standardized (mean=0, standard deviation=1) to allow for comparison across multiple datasets (Lundblad and Wright 2006).

Another habitat morphometric that was calculated was planar curvature, which calculates the rate of change in slope and measures the concavity or convexity of a surface (Pittman et al. 2011). Curvature is considered to be an important measurement of seafloor habitat due to the connection between seabed characteristics and current strength (Wilson et al. 2007). Vector Ruggedness Measure (VRM) quantifies substratum complexity by defining “ruggedness” as the variation in the three-dimensional orientation of pixels within a raster and has previously been used to delineate habitats of marine organisms (Beck 2000; Kostylev 2005). Orthogonal dispersion within a specific neighborhood is calculated to measure the three-dimensional orientation of grid cells

(Valentine et al. 2004). This results in ruggedness values ranging from zero (no terrain variation) to one (highest terrain variation).

Rockfish abundance does not generally exhibit an exact linear relationship with depth, displaying instead a parabolic relationship, with preferences being bracketed within a specific depth range (Love 1990; Stein 1992; Auster et al. 1995; Yoklavich and Greene 2000; Love and Yoklavich 2006; Love and York 2006). Thus, for each model that included depth as a predictor, a distinct model including the polynomial term Depth^2 was developed and compared using AIC scores (Zeileis et al. 2008).

Once the topographical values were computed for all habitat covariates, they were extracted to the transect point data using the Extract Multi Values to Points tool in ArcMap. The habitat covariates of interest are summarized in Table 1.

Analysis of video showed that Squarespot and Rosy rockfish were often encountered at the interface of rocky habitat and soft sediment, or “edge” habitat. For this reason, I decided to include distance to reef edge as an omnidirectional environmental covariate. I created a map of the study region classifying terrain as either “rock” or “soft sediment” by selecting a VRM cutoff value that adequately identified the region into the different substrate types, which delineated edge habitat. After comparing the original DEM to VRM rasters of varying cutoff values, it was determined that a VRM value of 0.002 was acceptable (Figure 3). These Euclidean distances between fish and reef edge were calculated within ArcMap by first converting the VRM raster (which indicates the rugosity of the substratum) to a polygon shapefile. Only those polygons containing VRM values greater than the established cutoff value for hard substrate were included in order to demarcate rocky reef from soft sediment. The polygons in this shapefile, indicating sections of rocky reef of explicit VRM values, were then merged so that individual rocky reefs were now represented by a single polygon. The shapefile containing these reef polygons was converted into polylines, and the distance to each line (representing reef edge) was acquired by joining the polyline shapefile to the shapefile containing the fish detection data using the Join tool.

Similarly, adult Rosy rockfish were observed preying periodically on Pygmy YOY. For this reason, distance to the nearest Rosy within a site was calculated for all Pygmy and Squarespot observations and included in the statistical analyses.

In order to achieve a balanced sampling design for analysis, non-detection points were randomly generated from the transect data equal to the number of species' detection points observed within a site (Bassett et al. 2018). Having an equal number of detection and non-detection points allowed me to compare areas of fish habitat usage without an overabundance of zeros in the data. To generate the non-detection points, fish detection points (geographic coordinates where fishes were observed along a transect) were first buffered by 5 m to eliminate the possibility of including potentially suitable habitat in the pool of locations available for selection. These are mobile species that exhibit relatively high site fidelity (Lowe et al. 2009), so a radius of 5 m around the coordinates of an observed fish was deemed appropriate; it was very likely that the habitat the fish was immediately associating with was very similar to the habitat type within a 5 m radius. Therefore, I did not want to potentially select this immediately adjacent habitat as a non-detection point. Once the detection points were buffered, the non-detection points were constrained to the transects sampled within a site, to represent where fishes were actually searched for and not observed. This process was repeated separately for the three age classes of each species, and then combined for a total of nine datasets for each species-site combination. For example, if 20 YOY Pygmy were encountered at Point Lobos, then 20 non-detection points were randomly generated along all transects sampled within Point Lobos, representing 20 locations where YOY Pygmy were searched for and not encountered. This dataset, containing a total of 40 points, was then combined with the Year-1 and Year-2 Pygmy datasets, resulting in a single dataset containing all the Pygmy data at Point Lobos. This balanced dataset was still considered to be zero-inflated by GLM standards, requiring me to utilize a different modeling approach than a traditional GLM.

Statistical Analysis

Hurdle models

Predictive models of species' distribution were developed using zero-adjusted negative binomial (ZANB) and zero-adjusted Poisson (ZAP) models, also commonly referred to as hurdle models (Zuur et al. 2009). These models are designed to handle excess zeros and overdispersion, and hence are useful for demersal fish count data (Santos et al. 2011; Turschwell et al. 2017). Most importantly, they have the ability to predict species' occupancy and abundance simultaneously from a single dataset, thus making them an ideal tool when handling count data.

The first part of the hurdle model predicts the probability of a species' presence via a binomial distribution given a set of environmental variables, while the second part models the species' abundance given that it is present at a location via a Poisson or negative binomial distribution. Not all environmental predictors need to be statistically significant in the "best" hurdle model; in many cases removal of non-significant variables results in a model with a poorer fit based on AIC and R^2 values. Furthermore, the predictor variables can be the same for the logistic regression and the count process of a hurdle model, but they often differ between the two procedures. This feature makes hurdle models very useful for describing distribution when a species' occupancy and abundance arise from distinct ecological processes. For example, consider a species that has been extirpated from a number of its previously inhabited habitats due to overfishing or land use development. In this case, the species' occurrence at a given site may be predicted by one variable such as historic trawling pressure, while its abundance when it is present is influenced by another set of environmental characteristics, such as suitable rock habitat.

The logistic regression in a ZAP or ZANB models the probability of presence versus absence, rather than the probability of a false zero versus other types of data. The equation for the zero-adjusted model is:

$$P(Y_i = y_i | X_i, z_i, \beta, \gamma) = \begin{cases} f_{zero}\left(0; z_i; \gamma\right), & if\ y_i = 0 \\ \left(1 - f_{zero}\left(0; z_i; \gamma\right)\right) \dfrac{f_{count}\left(y_i; x_i; \beta\right)}{1 - f_{count}\left(0; x_i; \beta\right)} & if\ y_i > 0 \end{cases} \quad (1)$$

Where y_i is the value of the dependent variable, z_i represents the number of predictors in the zero process, X_i represents the number of predictors in the count process, β is a vector of coefficients belonging to X, and γ is a vector of coefficients related to z. From Equation 1, the probability of measuring a zero count (f_{zero}) is modeled using a binomial distribution. The lower part of the equation states that the probability of measuring a non-zero value equals the probability that it is non- zero value multiplied with the probability determined by the count distribution (f_{count}). In other words, to measure a non-zero count, the environment must first cross the "hurdle" of producing a non-zero value. As illustrated by the denominator of the fraction in the lower portion of the equation, the count process excludes the probability of any zero counts. This is the reason these models are referred to as zero-truncated or zero-adjusted distributions.

Species' distribution models

Statistical analyses were accomplished using R statistical software and the package 'pscl' (Jackman 2010) and included data exploration protocols described by Zuur et al. (2010). Collinearity among predictor variables was investigated using correlation matrices and Variance Inflation Factors (VIF) (Montgomery and Peck 1992). If two variables were highly correlated (VIF>5), then only one was kept based on the likelihood of it being a predictor of rockfish abundance. This likelihood was determined after a literature review.

To assess model accuracy, each of the nine datasets were first partitioned in a 90/10 split whereby 90% of the data were randomly selected to be the "training data," while the remaining 10% were to be used as "test" data (Morrison et al. 2013). Akaike's

Information Criterion (AIC) was used to select the most parsimonious model from all possible combinations of covariates (Sakamoto et al. 1986), and validation of the optimal model was accomplished by examining residuals vs. fitted values (Anscombe and Tukey 1963). Model fit was assessed using a rootogram from the R package ‘countreg’ (Kleiber and Zeileis 2016). A rootogram illustrates a model’s accuracy via a histogram of the count values in a dataset, whereby the bar height represents the difference between the observed and expected count values. A bar hanging below the zero line on the y-axis indicates underfitting by the model for that particular count category, while a bar hovering above zero indicates overfitting. Ad hoc testing is required to determine the exact numerical difference between the two values, such as using the “predict” function in R to provide an estimated value of the count value of interest.

Each model was then used to predict the species’ count values given the site’s test data. To demonstrate that a model’s predictive power was linked to its spatial origins, Mean Absolute Error (MAE) and pseudo-coefficient of determination (R^2) values were recorded when the site-trained model was used to predict that site’s own test data, and when it predicted species’ distribution at the two other sites whose training data were not used to inform the model (Iampietro and Kvitek 2008). MAE measures the average magnitude of errors in a set of predictions without considering their direction. In other words, it is the average over the test sample of the absolute differences between the predictions and observations, where all individual differences have equal weight (Wilmott and Matsuura 2005). This can be represented by the equation:

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|x_i - x| \qquad (2)$$

Where n is the number of data points, x_i is the predicted measurement, and x is the observed value.

MAE was chosen as a more appropriate metric to evaluate model performance in this case rather than Root Mean Square Error (RMSE), because the error distribution was not Gaussian (Chai and Draxler 2014).

The coefficient of determination (R^2) is defined as the proportion of variance within the dependent variable that is explained by the independent variables. A pseudo-R^2 value has been developed for use with zero-adjusted models that avoids inflation due to the inclusion of irrelevant covariates in the model (Carlevaro and Hoareau 2012). This psuedo-R^2 value was utilized in this study in lieu of a traditional coefficient of determination using the R package “mhurdle”.

Spatial autocorrelation

Environmental variables are commonly correlated with one another, such that observations that are closer geographically would be more similar to each other than expected by random chance. This spatial autocorrelation (SA) may persist after accounting for measured environmental variables due to other non-measured spatial similarities or effects (Legendre 1993). Thus, observations made on transects that were nearer to each other may have naturally been more similar than observations made on transects separated by greater distances. This can result in Type I errors (false positives), which can reduce the explanatory power of variables in distribution models (Segurado et al. 2006). In other words, if SA is present in the response variable, the significance of the explanatory variables will be inflated.

To investigate lack of spatial independence in the data the residuals of each species' "global" model at each site were tested for significant SA using the R statistical package 'spdep' (Bivand and Wong 2018) using Moran's *I* coefficients and a binary weighted neighboring scheme. A species' global model included all habitat covariates as predictor variables without regards to significance. The global model did not incorporate a spatial component such as a random effects term. Moran's *I* is a correlation coefficient that measures how similar samples of a given variable are over varying spatial distances. Additionally, it calculates a p-value to evaluate significance. A Moran's *I* value that is close to zero indicates no spatial autocorrelation.

If significant SA was observed in the residuals of the global model, then four new models were constructed incorporating Transect as a random effect, in an attempt to mediate the effects of SA. Each of these four "spatial" models included the same habitat covariates as the global model, and utilized one of four different correlation structures (rational quadratic, exponential, spherical, and Gaussian; Pinheiro and Bates 2000). The first correlation structure represents a rational quadratic spatial correlation structure and is illustrated by the equation:

$$1 \,/\, (1 + (r \,/\, d)^2) \qquad (3)$$

Where d denotes the range of points in three-dimensional space, and r is the distance between two observations.

An exponential spatial correlation structure measures the correlation between two observations at a distance r apart and is denoted by:

$$exp\left(-(r / d)\right) \tag{4}$$

A spherical spatial correlation structure between two observations at distance $r < d$ apart is illustrated as:

$$1 - 1.5(r / d) + 0.5(r / d)^3 \tag{5}$$

Finally, a Gaussian spatial correlation structure was constructed, where the correlation between two observations at r distance apart is denoted by:

$$exp(-(r / d)^2) \tag{6}$$

The normalized residuals of these four spatial models were tested for significant SA using a binary weighted neighboring scheme (Ord and Getis 1995). The results of these tests were compared to the results of the global models to determine whether the addition of Transect as a random effect and the inclusion of a spatial correlation structure removed the spatial autocorrelation in the residuals, and if so, which spatial correlation structure was most appropriate to utilize. If the spatial models did not resolve the spatial autocorrelation in the residuals, analysis proceeded with the simple model that did not incorporate a random effects term; in these instances, there were some spatial processes or effects at play that could not be accounted for statistically that were potentially affecting model results. However, multiple studies have utilized non-spatial models after investigating the effects of SA; these studies similarly concluded that either the incorporation of a random effects term did not reduce the effects of SA, or the resulting spatial models were less interpretable and not as useful in predictive modeling (Young et al. 2010, Wedding and Yoklavich 2015).

RESULTS

Spatial Autocorrelation

There was significant spatial autocorrelation in the data from Pt. Lobos and Pt. Sur based on the Moran's *I* values of the residuals of the global models. When Transect was incorporated as a random effect into the spatial models, the Moran's *I* values of the residuals at Pt. Lobos did not improve. Similarly, the residuals from the global model developed for the Pt. Sur data indicated there was significant spatial autocorrelation in the data. The SA at Pt. Sur was improved by the addition of a spatial component, but this addition did not remove the significant SA entirely. No significant SA was detected at Portuguese Ledge after examination of the residuals of the global model (Table 5). Because the SA found at Pt. Lobos and Pt. Sur could not be resolved by the spatial models, they were not used for interpretation. Analyses continued on the simple, non-mixed models that did not incorporate a spatial component.

General Site Model Results

Covariates that were highly correlated were removed from analysis. The habitat variable b-BPI was found to be highly correlated with f-BPI and depth (VIF= 15.69) and was therefore excluded from further analyses. For each model that included depth as a predictor, a distinct model including Depth^2 was developed and the AIC scores of the two models were compared. Although a parabolic relationship was evident in the Depth^2 data, including the polynomial term did not significantly improve any of the models. Overall,

all three species were found to be statistically associated with relief type, sloping terrain, and cardinal direction of slope faces. However, the degree to which these characteristics affected fish associations varied among study locations.

Of the three study sites, Pt. Lobos had the greatest extent of available high relief rocky substrate. This high complexity habitat supported the greatest number of Squarespot and Pygmy rockfish at all three study sites (Table 2). Squarespots at Pt. Lobos had a significant positive relationship with high relief, and a significant negative relationship with mixed substrata (Table 3). Pygmys however, continued to demonstrate a negative relationship with high relief and mixed relief, with only a near-significant positive relationship to moderate complexity hard substrata. Pygmy rockfish did not appear to have a direct association with relief, instead showing more generalist habitat distributions. At Pt. Lobos and Pt. Sur where high complexity rock was abundantly available, they demonstrated a significant negative relationship to the rugose reefs. In none of the Pygmy models does any relief category appear to have a significant positive effect on occurrence or abundance. Instead, Pygmy abundances seem to be related to other environmental predictors such as aspect, depth, and BPI.

The Rosy rockfish at Pt. Lobos exhibited a significant positive relationship with the two environmental covariates correlated with rugosity, VRM and BPI (Table 3). Furthermore, this was the only site at which these two variables were significant predictors in the Rosy rockfish models. This suggests that when high relief habitat is readily available, we can predict Rosy rockfish distribution based on the emergent high relief reefs. However, where rugose structure is not as readily available, as is the case at

Portuguese Ledge, this species will alter their habitat preferences and other environmental variables become better at predicting the species' distribution. At Portuguese Ledge and Pt. Sur, VRM and BPI were not included in the best-fitting model; instead, Rosy rockfish at this site were found to associate significantly with slope, indicating steeper sloping terrain may be the best habitat available to these fish when rugose rocky reefs are scarce.

Modeling of Squarespot Rockfish

Overall, there were 766 observations totaling 1,319 Squarespot individuals, comprised of 136 observations from Portuguese Ledge, 446 observations from Point Lobos, and 184 observations from Point Sur (Table 2). An observation can be one instance of any number of fish observed within a one-second interval (such as a school), whereas individual counts represent an individual fish. Squarespot were found over all habitat types, but probability of occurrence was highest over boulder and cobble substrate. This species consistently associated positively with high relief rock, and was observed over soft sediment at Portuguese Ledge. Squarespot rockfish were also more likely to be encountered on slopes facing into the direction of the prevailing California current (north to northeast).

Portuguese Ledge

The best hurdle model predicting Squarespot distribution at Portuguese Ledge included depth, relief, and aspect (Table 3). This model accounted for approximately 24% of the variability in Squarespot distribution (Table 4).

Examining response curves of individual predictors, the probability of Squarespot presence was greatest at locations with substrata in the Mhard relief category. The probability of Squarespot presence was significantly influenced by aspect and decreased with depth from 70 to 100 m (Figure 4). Squarespot abundance decreased with depth at Portuguese Ledge (Figure 5).

Point Lobos

The best hurdle model predicting Squarespot distribution at Point Lobos included distance to reef edge, relief, and distance to nearest Rosy rockfish. This model accounted for approximately 5% of the variability seen in this species' distribution at this site. The probability of Squarespot occupancy at a site was highest at locations with moderate hard relief, and declined with increasing distance to the nearest Rosy (Figure 6). Distance to reef edge was the only significant predictor of Squarespot abundance at Point Lobos, where Squarespot abundance decreased with greater distances to a rocky reef outcropping (Figure 7).

Point Sur

The best hurdle model predicting Squarespot distribution at Point Sur included distance to reef edge, relief, aspect, VRM, and distance to nearest Rosy, and accounted for approximately 38% of the variance in distribution at this site. As was the case at the other sites, Squarespot presence was most probable in areas of moderate hard relief (boulders and cobble). Probability of occurrence declined with increasing aspect, higher rugosity (larger VRM values), and distance to nearest Rosy rockfish (Figure 8). Similar to the pattern observed at Point Lobos, Squarespot abundance at Point Sur declined as distance to reef edge increased (Figure 9).

Modeling of Rosy Rockfish

A total of 592 observations of 631 Rosy individuals were collected from the three study sites: 151 observations from Portuguese Ledge, 91 observations from Point Lobos, and 350 observations from Point Sur. In each of the hurdle models developed for all three sites, the Rosy rockfish data fitted a Poisson distribution and abundance was not significantly influenced by any of the habitat covariates.

Portuguese Ledge

The best hurdle model predicting Rosy occurrence at Portuguese Ledge included relief and slope. This model accounted for approximately 17% of the variability in the species' distribution at Portuguese Ledge. Like Squarespot, the probability of Rosy rockfish presence was highest at locations with boulder or cobble (Mhard relief) substrata (Figure 10). Probability of occurrence slightly increased with steeper slopes at Portuguese Ledge.

Point Lobos

The selected model for predicting Rosy occurrence at Point Lobos included relief, aspect, VRM, and f-BPI, and accounted for approximately 13% of the variability seen in the distribution of Rosy at Point Lobos. However, unlike the Rosy population found at Portuguese Ledge, the probability of occurrence was lowest at locations comprised of substrata of the Mhard complexity (Figure 11). Probability of occurrence also decreased with aspect, and increased with increasing VRM and f-BPI values.

Point Sur

The best hurdle model predicting Rosy rockfish at Point Sur included patch length, relief, depth, and distance to reef edge, and it accounted for approximately 20% of the variability in Rosy distribution. The relief category Mhard was the best predictor of Rosy occurrence at this site, while the probability of Rosy presence declined with increasing patch length, depth, and distance to reef edge (Figure 12).

Modeling of Pygmy Rockfish

Pygmy rockfish were the most abundant species surveyed; 915 Pygmy observations were made totaling 1,786 individuals from the three study sites. There were 278 observations from Portuguese Ledge, 510 observations from Point Lobos, and 127 from Point Sur. This species was the only one of the three to associate negatively with Lhard. Similar to the Squarespot rockfish, Pygmys were more likely to be found over current-facing slopes.

Portuguese Ledge

The best model predicting Pygmy distribution at Portuguese Ledge were relief, patch length, aspect, and f-BPI, wherein patch length and aspect predicted the logistic regression portion of the model, and relief, patch length, aspect, and f-BPI modeled the species’ abundance. This model accounted for over 61% of the variability found in Pygmy distribution at Portuguese Ledge. The probability of occurrence increased with longer patch lengths and decreased with increasing aspect (Figure 13). Pygmy abundance was greatest in areas of highest relief (Lhard). Abundance also increased with aspect, patch length, and greater f-BPI values, corresponding to higher terrain elevation (Figure 14).

Point Lobos

The model that best described Pygmy distribution at Point Lobos included depth, relief, and distance to nearest Rosy rockfish. All three variables predicted Pygmy occupancy, while only depth and distance to nearest Rosy rockfish modeled Pygmy abundance given its presence. This model accounted for little more than 1% of the variability in Pygmy distribution, but was confirmed to be the best-fitting hurdle model following established model selection protocols. Mhard was again the best predictor of occupancy; this probability increased with increasing depth and distance to nearest Rosy (Figure 15). At Point Lobos Pygmy abundance declined with increasing depth and distance to Rosy (Figure 16).

Point Sur

The selected model to predict Pygmy distribution at Point Sur included relief, depth, distance to reef edge, distance to nearest Rosy, and curvature. This model accounted for nearly 60% of the variability found in the species' distribution at this site. Probability of occupancy at Point Sur was highest in areas with moderate rocky relief (Mhard relief category), and increased sharply with increasing depth (Figure 17). Contrary to what was expected, the probability of occupancy declined with Rosy rockfish distance and increased slightly with increasing distance to reef edge. Abundance however decreased with increasing distances to Rosy rockfish, and increased with greater curvature (i.e. surfaces of higher concavity) (Figure 18).

Comparison of Site-Trained Models

Species' distribution, predicted using training data, demonstrated the highest accuracy was achieved when using models developed at the site where the test dataset originated. Utilizing a model developed at a different site to predict species' distribution resulted in lower R^2 values (Table 4). For instance, when predicting Pygmy distribution at Portuguese Ledge, the model trained by data gathered from Portuguese Ledge had an R^2 value of 0.61. When the same population's distribution was predicted using the model informed by the species' patterns observed at Point Lobos, the R^2 value was much lower (0.05), indicating poor model fit. Alternatively, when the model developed at Portuguese Ledge was used to predict Pygmy distribution at Point Lobos, the R^2 value was approximately 0.0002. Abundance of all three study species was far lower at Portuguese Ledge where rocky habitat is rare. Individual habitat associations at this site were different than those observed at Pt. Lobos and Pt. Sur; the distributional models developed at these two high-relief sites did not accurately predict the species' distributions at Portuguese Ledge.

With the exception of the model trained at Point Sur to predict Rosy (Sur.r) and the model trained at Point Lobos to predict Pygmy (Lobos.w) distribution, the models trained to their specific sites had lower MAE values than when models trained at one of the two other sites were used to predict species distribution. When comparing models, a lower MAE value indicates a better fit. As expected, most of the models trained and tested at the same site had the lowest MAE as compared to those models trained at one site and used to predict test data from another. This low error rate can be visualized in the rootograms of the models (Figures 19-21).

DISCUSSION

This study was the first to investigate adaptive habitat usage of rockfish species across a large swath of their geographic range. The utilization of standard habitat characterization protocols allowed me to compare patterns of habitat associations of commonly studied demersal fish species. This study was successful in demonstrating that a species' habitat usage varies in accordance to the relative physical surroundings. Habitat factors that may have been important in predicting occurrence or abundance in one area were not always found to be important in another. Every species' distribution model developed at a distinct site resulted in a unique set of variables predicting the species' distribution.

Relief was an important predictor of species' occupancy in nearly every model developed, with the exception of Pygmy distribution at Portuguese Ledge, where it was only a predictor of abundance. Pygmy rockfish were more often associated with substrata in the Mhard relief category, rather than the more structurally complex rock pinnacles and ledges. This coincides with the findings of Tissot et al. (2007) for Pygmy rockfish at Heceta Bank, Oregon, which found that boulder and cobble habitat supported high Pygmy rockfish densities. Contrary to the conclusions of many previous studies however (Love and Yoklavich 2006; Love et al. 2009; Young 2010), Rosy rockfish in this study were not associating strongly with areas of highest relief (Lhard). Instead, the areas with the highest probability of occurrence for all three rockfish species consisted of substrate in the moderate hard relief category (Mhard), which includes boulders and cobble. This result was similar to the findings of Wrubel (2013) along the Olympic coast of Washington.

As expected, occurrence and abundance for all three species of rockfish were lowest over soft substrates. However, some populations utilized these habitats when rocky relief was rare. For example, at the Point Lobos and Point Sur study sites no individuals of Squarespot were found over soft substrate, which is expected given what we know about this species' preferred habitats. However, at Portuguese Ledge, where soft sediment accounted for over 22% of the substrate surveyed, several individuals were observed over mud or sand (n=7). At Point Lobos and Point Sur, soft sediment made up only 3% and 6% of the total substrate surveyed across all transects respectively, and there were no observations of Squarespot or Pygmy rockfish on soft sediment at these sites, whereas 19 Pygmy individuals were observed over sand or mud at Portuguese Ledge. These results indicate that if their preferred habitat is not available, rockfishes may alter their preferences and shift to less preferred substrata. These shifts in habitat use and their effects on population density should be taken into consideration when distributional models are being applied by resource managers, specifically in the planning of EFH and MPAs.

Species' responses to mixed relief was highly variable. The population of Squarespot at Point Lobos, for example, were less likely to associate with mixed substrata, while at Point Sur their abundances over this relief type were higher in comparison. Mixed relief was not found to be significant in any of the three models developed for the Rosy rockfish and had a significant negative effect on Pygmy occupancy relative to the three other relief types only at Point Lobos. Further subdividing the mixed category of substrata may shed some answers as to how these species are specifically associating with these areas of integrated relief. Point Sur had the highest abundance of Rosy rockfish of all three study sites and had slightly less overall high relief substrata. Given Rosy rockfish predation on YOY Pygmys, it may be that Pygmy rockfish at Pt. Sur are being displaced from their preferred habitat among the high relief rock. This may explain why more Pygmys were observed over mixed substrata at Pt. Sur than anywhere else.

Distance to reef edge was a significant predictor in four of nine of the species' distribution models. Overall, most of the rockfish species studied followed a trend of decreasing abundance with increasing distance to the edge of a rocky reef; the farther from a rocky reef, the less likely that species was encountered. Because these species are known to associate strongly with high relief rock (Love et al. 2002), this result concurs with established hypotheses regarding rockfish habitat associations. One exception to this pattern were the Pygmy rockfish at Pt. Sur, whose probability of occurrence actually increased farther from the reef edge. Directional distance was not quantified in these data, so it cannot be stated with certainty whether the Pygmys were found farther away from reefs over soft substrate, or if large numbers of Pygmys were observed near the centers of large reefs, the theoretical distances being the same. The probability of Pygmy occurrence was highest over substrata of the Mhard relief category, so the latter instance is the most likely scenario. In any case, Pygmy rockfish at Pt. Sur did not associate as strongly with edge habitat as the other two target species, suggesting this metric is not a good predictor for this species. Boulders accumulate around the bases of high relief rocky reefs and provide shelter and foraging space for piscivores such as Rosy rockfish which may be one explanation for the positive association observed between the two more aggressive rockfish species and edge habitat.

The probability of Squarespot occupancy decreased with increasing distances to the nearest Rosy rockfish. This result is reasonable when considering the fact that Rosy and Squarespots occupy similar habitats (Love et al. 2002; Love and Yoklavich 2006; Iampietro et al. 2008), and so this negative trend may be suggestive of a surrogate predictor; Rosy rockfish presence in an area may be indicative of good habitat or favorable conditions, from which Squarespot rockfish are likewise benefitting from. Pygmy rockfish demonstrated a more complicated relationship with their nearest Rosy neighbor; the probability of Pygmy occurrence increased with greater distances to Rosy rockfish. However, Pygmy abundance significantly decreased with increasing distances to Rosy rockfish, suggesting they are potentially selecting alternative habitats to Rosy rockfish (which coincides with the results detailing Pygmys' relationship with edge habitat). Rosy rockfish may be competing with or preying on Pygmys along the favorable edge habitat, forcing Pygmy rockfish to seek shelter either further into the reef, or out into the soft or mixed substrata.

The potential role that benthic invertebrates play in influencing rockfish habitat associations was not investigated in this study. However, including invertebrate covariates such as percent cover or species diversity may have improved the models' predictive ability, particularly when the inclusion of abiotic factors alone did not yield an accurate, well-fitting model. This region of the central California coast does not harbor a significant population of the large structure-forming benthic invertebrates commonly thought to provide fishes with biogenic habitat, such as members of the genus *Antipathes* (black corals), barrel sponges (Class Demospongiae), or taxa in the Order Alcyonacea (soft corals) (Tissot et al. 2006; Wrubel 2013; Huff et al. 2013; Whitmire et al. 2017). However, groups of fish were observed over large swaths of brachiopods (Order Terebratulida), sea pens (*Ptilosarcus gurneryi*), and sea whips (*Halipteris willemoesi*) which form their dense aggregating beds in soft mud or sand, substrata with which rockfish do not typically associate. These observations suggest that large aggregations of smaller invertebrates may provide alternative structure for rockfish in the absence of high relief rock. Comparing rockfish densities between areas of hard substrate, soft substrate, and soft substrates with large densities of invertebrates may yield interesting results, particularly in sites that have a limited availability of hard substrata such as Portuguese Ledge. These invertebrate aggregations may be vital in providing habitat and structure in places that are bereft of high relief rock, where large abundances of rockfishes may not otherwise be able to persist.

The application of hurdle models proved to be a suitable method to predict species' abundances at distinct sites within the central California coastal region. Most of the zero-adjusted models developed demonstrated adequate fit, and predicted the count distributions from each site's test data reasonably well (R^2 values ranged from 0.2 – 0.6 in six out of nine models). When these site-trained models were applied to differing sites however, the resulting fit was consistently poorer. This supports the idea that differences in regional physical and ecological conditions play a key role in structuring fish assemblages and determining how a species interacts with habitat, which may potentially affect the way managers implement predictive distribution models. Site-level variation in terrain and availability of high relief rocky reef on the order of kilometers may be responsible for major changes in fish behavior and ecology. From the results of this study, it appears the availability of rugose structure and sloping terrain not only alters how rockfishes interact with habitat, but also impacts the local population's overall abundance.

The site-trained models that yielded low pseudo-R^2 values generally had the highest percentage of YOY and Year-1 individuals at a site as compared to the species' models with better fits. This poor fit may be attributed to the fact that juvenile rockfish behave and interact with habitat differently than their adult counterparts (Lenarz et al 1991; Larson et al. 1994; Johnson et al. 2001; Lindholm et al. 2003; Laidig et al. 2009; Hallenbeck et al. 2012). For example, YOY Pygmy rockfish exhibit strong schooling behavior, and were often observed schooling over soft substrate in deep canyons formed between large rock reefs. Adult Pygmy rockfish however, were more often associated with higher relief, in particular cobble and boulder habitat in the Mhard relief category. Grouping a species without respect to developmental stage can result in convoluted and inaccurate distributional models, due to the differences in habitat usage exhibited by these fish at different stages in their life history. This is why initially the data on these fishes were collected by age class so that they could be modeled separately. However at most of the sites, analyzing the data by age class resulted in small sample sizes that could not be modeled using such complicated techniques.

Hurdle models have the potential to serve as powerful tools in multivariate predictive modeling. In many systems, species’ distributions of occupancy and abundance arise from distinct ecological processes that can vary on spatial and temporal scales. This is why in many cases, it is difficult for researchers to accurately model certain species’ distributions, or why they seem to be able to develop an accurate distributional model in one area but not another. The hurdle models developed in this study were most accurate when modeling the distribution of aggregating species (groups comprised of three to nine individuals) that demonstrate high site fidelity. Rosy rockfish, which are commonly observed as solitary individuals (Anderson et al. 2009), did not conform well to the parameters of the zero-adjusted distributions. This is evidenced by the fact that the Rosy models had a lower average R^2 value than the two other species modeled, and that no environmental variables were significant predictors of Rosy abundance perhaps because abundance data were often binomial. There was only one instance of more than two Rosy rockfish observed together at any of the three study sites. It may be more appropriate to model solitary species such as Rosy rockfish in terms of presence/absence data, rather than count or density data. For aggregating or schooling species however, these models proved to be relatively accurate. If the goal is to conserve vulnerable populations, then we must ensure we are not going to waste resources by using models that may be misinformed. This study demonstrates that a robust knowledge of the differences in regional physical and ecological composition is essential when applying distributional models across a broad spatial scale, as we are seeing more and more evidence of adaptive habitat usage by rockfish species. If rockfishes' habitat associations are not static but rather subject to change dependent upon regional factors, managers must take such factors into account and attempt to predict their effects on rockfish population densities. Doing so will require more flexible spatial planning strategies, but may protect vulnerable stocks better than current strategies relying on broad generalizations can.

TABLES

Table 1. Summary of the habitat covariates used in statistical analyses.

Variable	Definition
Depth	Distance in meters from the surface.
$Depth^2$	Polynomial term used to compare models including Depth as a significant predictor
Relief	Grouping based on complexity of substrata. Lhard (highest relief), Mhard, Soft (lowest relief), and Mix.
Slope	The maximum rate of change in elevation between pixels. Measured in degrees.
Aspect	Slope orientation measured in continuous degrees from 0-360, with indices of cardinal direction. A flat surface is indicated by a -1.
Curvature	Calculates the rate of change in slope and measures concavity or convexity of a surface.
Bathymetric Positioning Index (BPI)	Characterizes features in a bathymetric DEM as positive or negative in relation to surrounding features. Calculated by comparing the elevation of a cell with the mean elevation of surrounding cells. Analyzed at fine-scale (scale factor of 36) and broad-scale (scale factor of 200).
Vector Ruggedness Measure (VRM)	Measures changes in slope and aspect simultaneously between pixels to represent terrain complexity. Ranges from 0-1.
Distance to Reef Edge	Euclidean distance from observed fish or non-detection point to nearest rocky reef edge.
Distance to Rosy	Euclidean distance from observed fish detection or non-detection point to nearest observed Rosy rockfish within a site.

Table 2. Summary of sample sizes of species at each of the study sites.

Site	Dives	Transects	Squarespot	Rosy	Pygmy	Total
Portuguese Ledge	5	13	184	112	965	1,261
Pt. Lobos	7	15	693	135	861	1,689
Pt. Sur	8	17	257	384	243	884

Table 3. Model covariates significant to the logistic regression portion of the selected hurdle model, predicting species' occurrence at a given location. Column names indicate the species (s = squarespot, r = rosy, p = pygmy) and study site (PD = Portuguese Ledge, PL = Pt. Lobos, PS = Pt. Sur). A plus sign (+) indicates a significant positive effect by the variable, while a minus sign (-) indicates a significant negative effect. A zero (0) indicates that although the specified covariate was included in the best-fitting model, its effects on the model were not statistically significant. DRE=Distance to Reef Edge, DTR=Distance to Rosy.

Covariate	s.PD	s.PL	s.PS	r.PD	r.PL	r.PS	p.PD	p.PL	p.PS
Lhard	+	+	+	**0**	**0**	+		-	-
Mhard	+	**0**	+	**0**	-	+		**0**	**0**
Mix	**0**	-	+	**0**	**0**	**0**		-	**0**
Soft	**0**			**0**	**0**	-			
Aspect	-		-		-		-		
Depth	-					-		+	+
VRM			**0**		+				
Slope				+					
f-BPI					+				
Patch Length						-	+		
Curvature									
DRE						**0**			**0**
DTR		-	-					+	-

Table 4. Model covariates significant to the count process of the selected hurdle model, predicting species' abundance at a given location. Note no variables were predictors of Rosy rockfish abundance.

Covariate	s.PD	s.PL	s.PS	p.PD	p.PL	p.PS
Lhard				-		
Mhard				**0**		
Mix				**0**		
Soft				**0**		
Aspect				+		
Depth	-				-	
VRM						
Slope						
f-BPI				+		
Patch Length				+		
Curvature						**0**
DRE		**0**	-			
DTR					-	-

Table 5. Pseudo-coefficient of determination and Mean Absolute Error values of each Squarespot model. Columns contain the name of each study site whose dataset is being predicted by each model. Models in each row are named by site and species abbreviation. Highlighted in green are the optimal values for each metric that coincided with what was expected. Blue indicates optimal values that did not coincide with expected results.

	Ledge R^2	Lobos R^2	Sur R^2		Ledge MAE	Lobos MAE	Sur MAE
Ledge.s	0.246	0.0231	0.0321	Ledge.s	0.481	4.04	6.87
Lobos.s	0.00328	0.0583	0.00797	Lobos.s	0.675	0.824	0.725
Sur.s	0.000256	0.0459	0.437	Sur.s	0.657	0.731	0.356

Table 6. Pseudo-coefficient of determination and Mean Absolute Error values of each Rosy model.

	Ledge R^2	Lobos R^2	Sur R^2		Ledge MAE	Lobos MAE	Sur MAE
Ledge.r	0.207	0.0617	0.0279	Ledge.r	0.438	0.473	0.522
Lobos.r	0.000417	0.128	0.062	Lobos.r	0.507	0.427	0.593
Sur.r	0.0163	0.0445	0.202	Sur.r	0.494	0.552	0.471

Table 7. Pseudo-coefficient of determination and Mean Absolute Error values of each Pygmy model.

	Ledge R^2	Lobos R^2	Sur R^2		Ledge MAE	Lobos MAE	Sur MAE
Ledge.p	0.614	0.000223	0.00404	Ledge.p	2.15	2.99	1.09
Lobos.p	0.0513	0.0119	0.00279	Lobos.p	1.58	1.01	3.77
Sur.p	0.00988	0.0089	0.592	Sur.p	2.74	2.49	0.565

Table 8. Moran's *I* values of the residuals from the global model of Portuguese Ledge ("orig"), and the subsequent mixed models incorporating Transect as a random effect as the spatial component. The abbreviated names of each model refer to the spatial model's correlation matrix. The statistical significance threshold is 0.05.

	observed	expected	sd	p-value
orig	0.0512	-0.00366	1.11	0.133
sph	0.0342	-0.00366	0.768	0.221
gaus	0.0377	-0.00366	0.84	0.2
rat	0.0412	-0.00366	0.926	0.177
exp	0.0381	-0.00366	0.848	0.198

Table 9. Moran's *I* values of the residuals from the global model of Pt. Lobos ("orig"), and the subsequent mixed models incorporating Transect as a random effect as the spatial component.

	observed	expected	sd	p-value
orig	0.362	-0.0013	13.1	<0.001
sph	0.354	-0.0013	12.8	<0.001
gaus	0.258	-0.0013	9.77	<0.001
rat	0.265	-0.0013	9.89	<0.001
exp	0.323	-0.0013	11.7	<0.001

Table 10. Moran's *I* values of the residuals from the global model of Pt. Sur ("orig"), and the subsequent mixed models incorporating Transect as a random effect as the spatial component.

	observed	expected	sd	p-value
orig	0.185	-0.0028	3.78	<0.001
sph	0.17	-0.0028	3.56	0.00019
gaus	0.145	-0.0028	3.03	0.00122
rat	0.142	-0.0028	2.96	0.00157
exp	0.145	-0.0028	3.017	0.00128

FIGURES

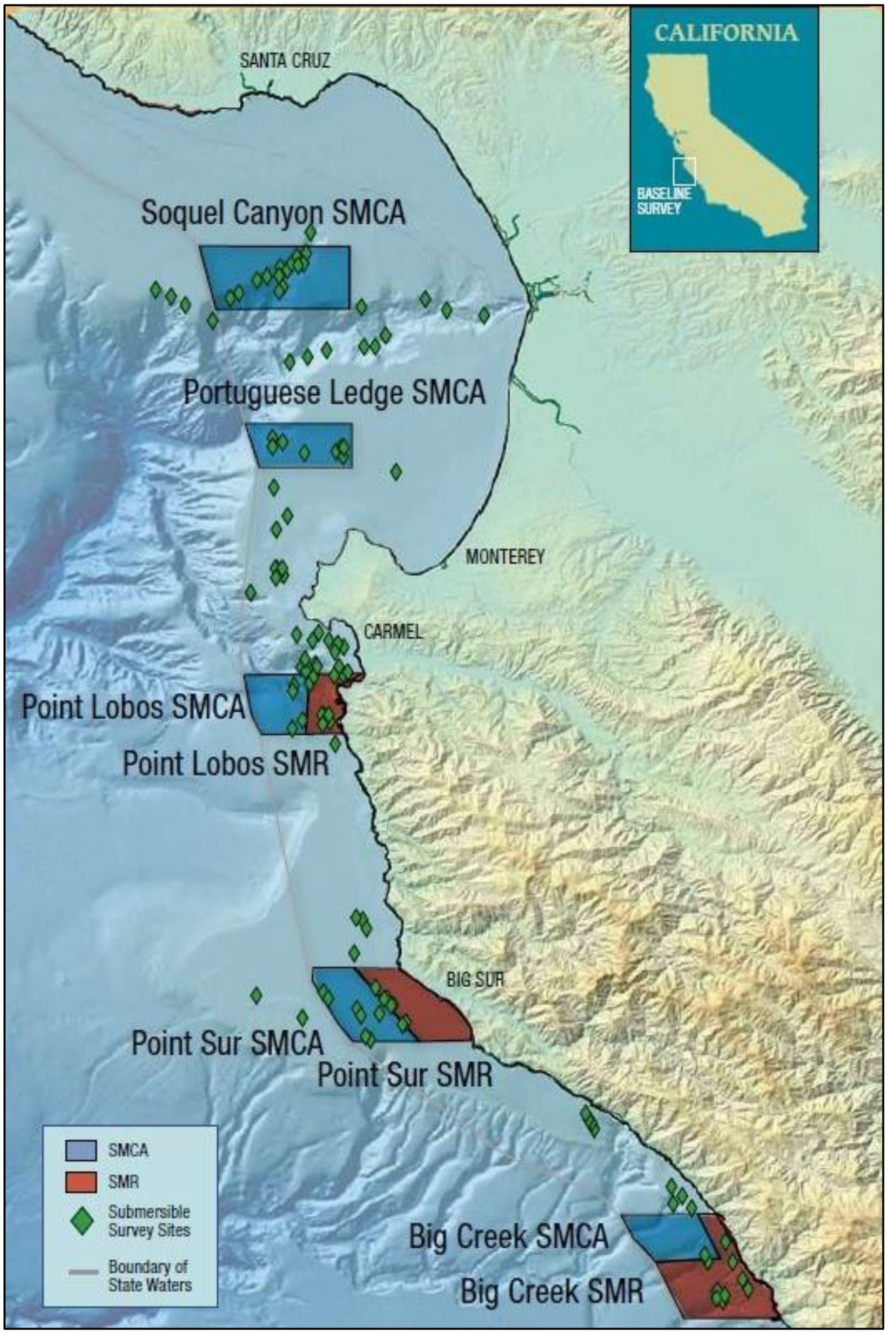

Figure 1. Map of 2007 deep-water submersible dives from the baseline surveys of MPAs and reference sites off the central coast of California. Reprinted from Starr et al. (2008).

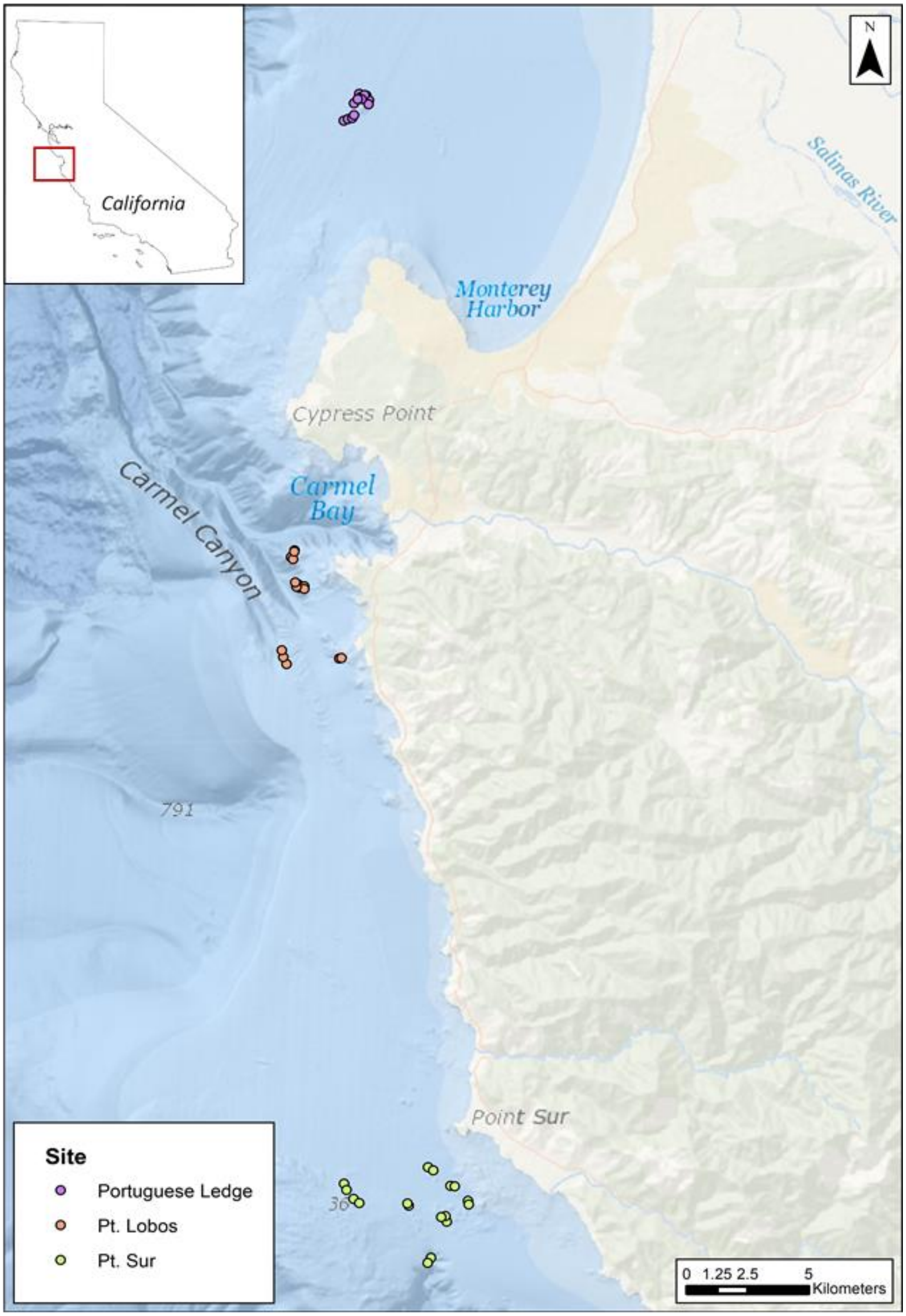

Figure 2. Map of transects selected from *Delta* surveys included in this study. Each location indicates the starting geographic coordinates of a transect.

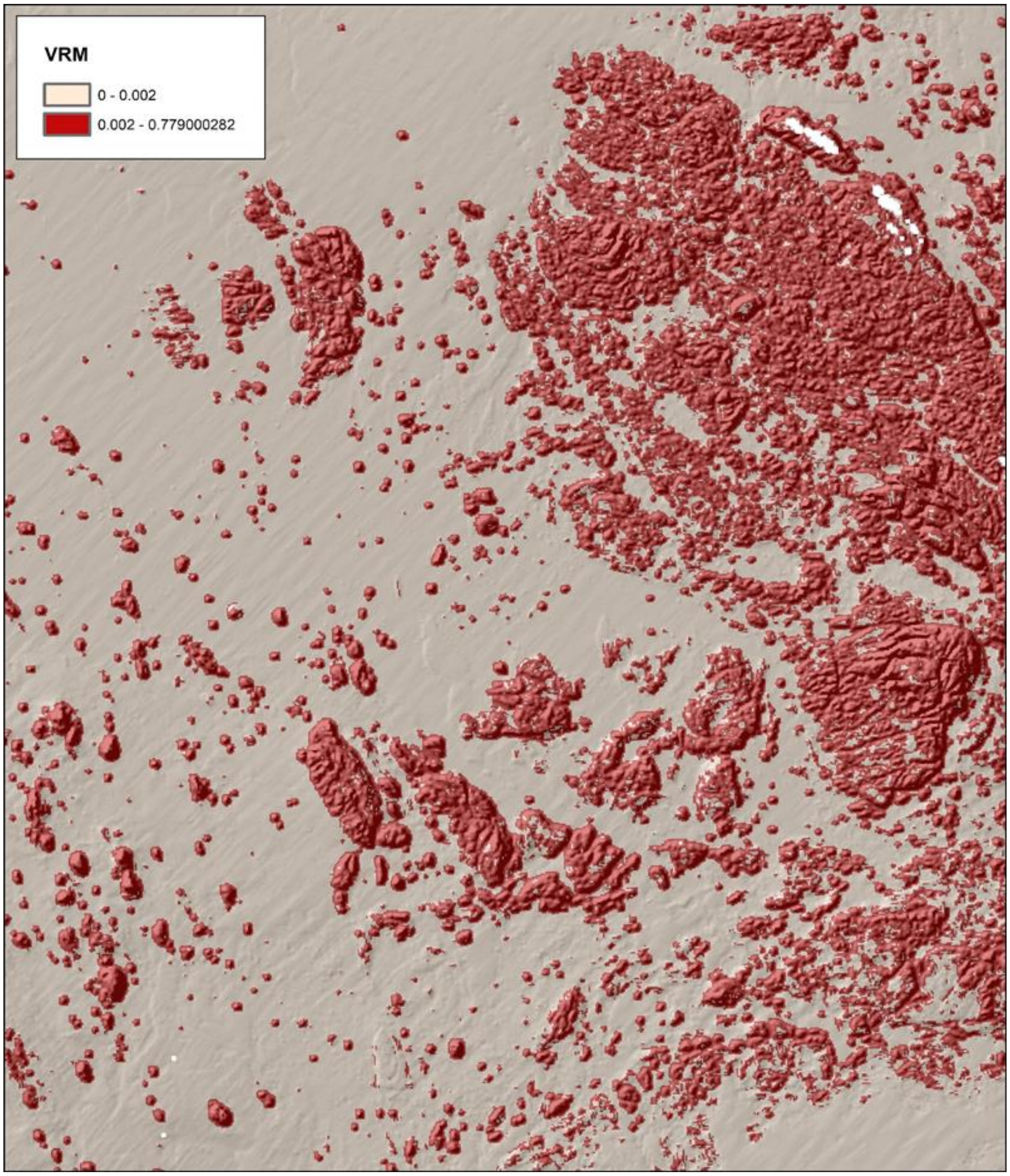

Figure 3. Illustration of Vector Ruggedness Measure (VRM) derived from the bathymetric DEM at the Point Sur study site.

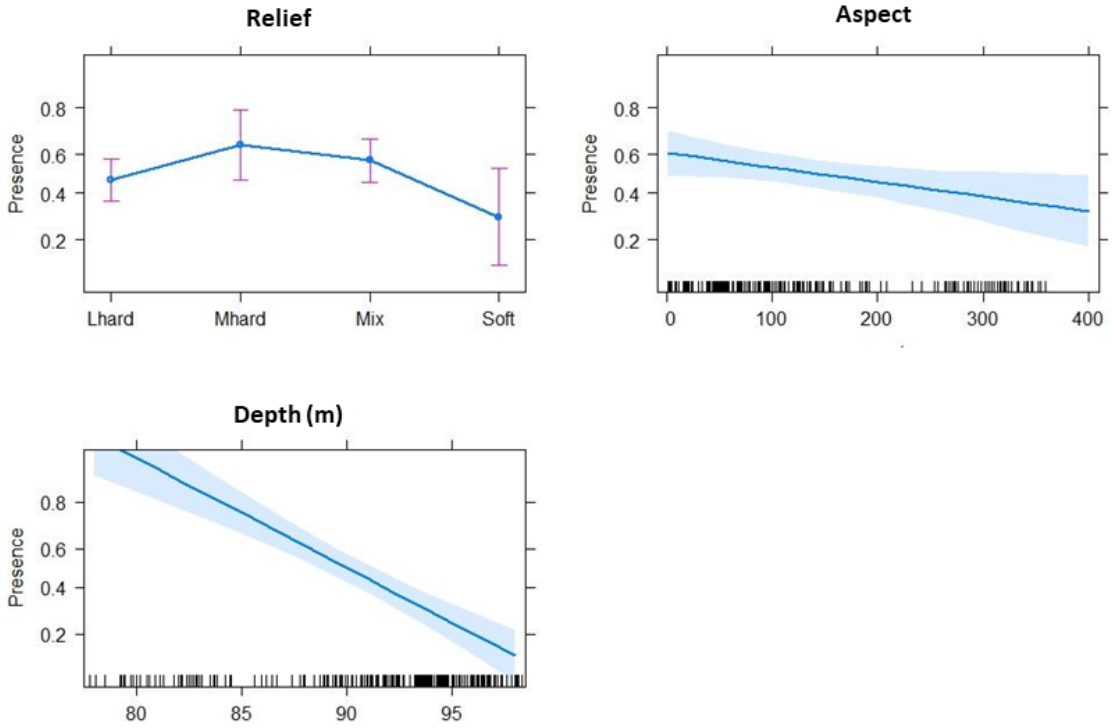

Figure 4. Effects of covariates on Squarespot occupancy at Portuguese Ledge. The vertical axes are labelled on the probability scale, and a 95% confidence interval has been drawn around the estimated effect.

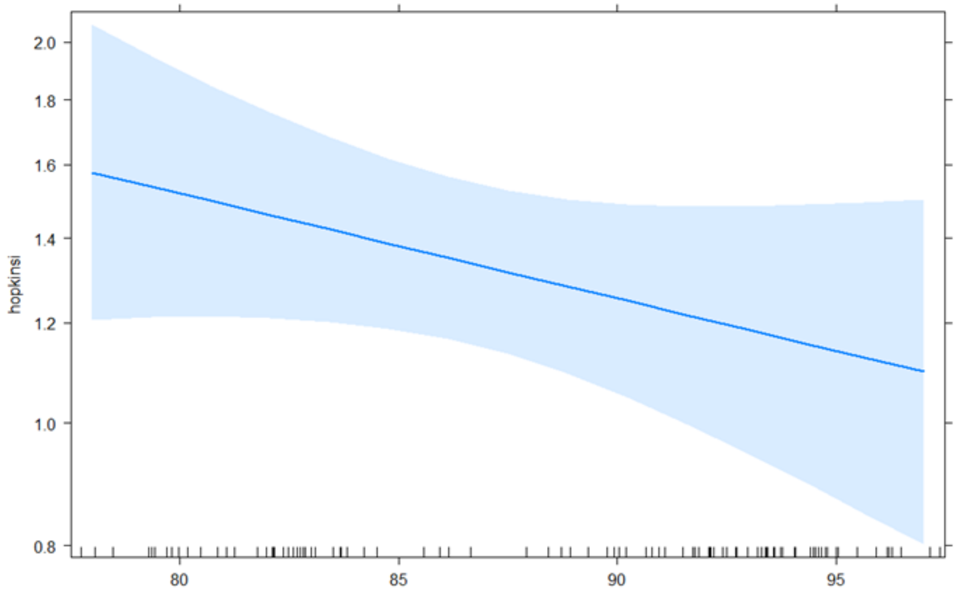

Figure 5. Effect of depth on Squarespot abundance at Portuguese Ledge.

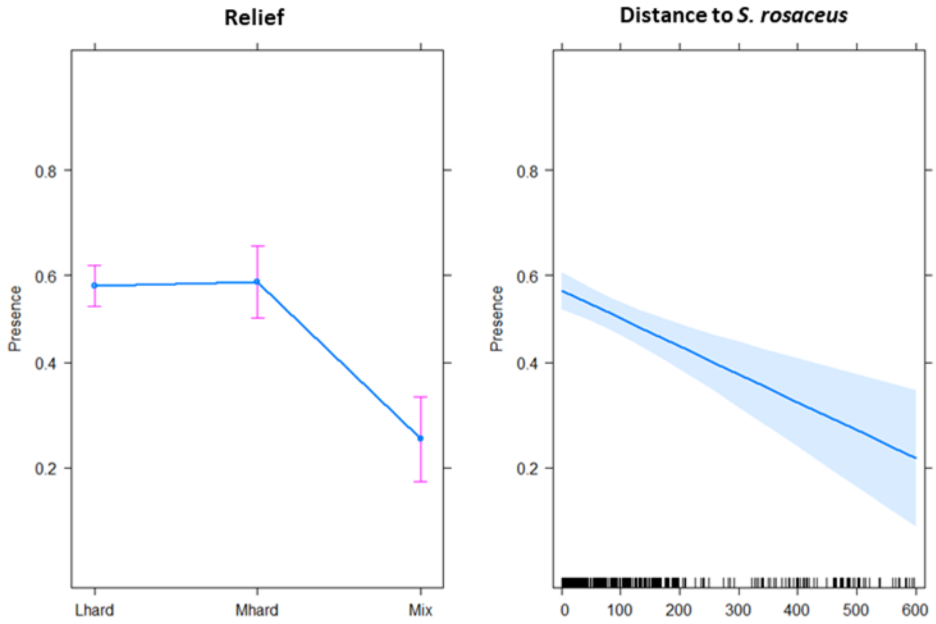

Figure 6. Effects plot of relief and distance to nearest Rosy rockfish on Squarespot presence at Point Lobos. No observations were found, and no non-detection points were generated over substrata classified as “Soft” relief.

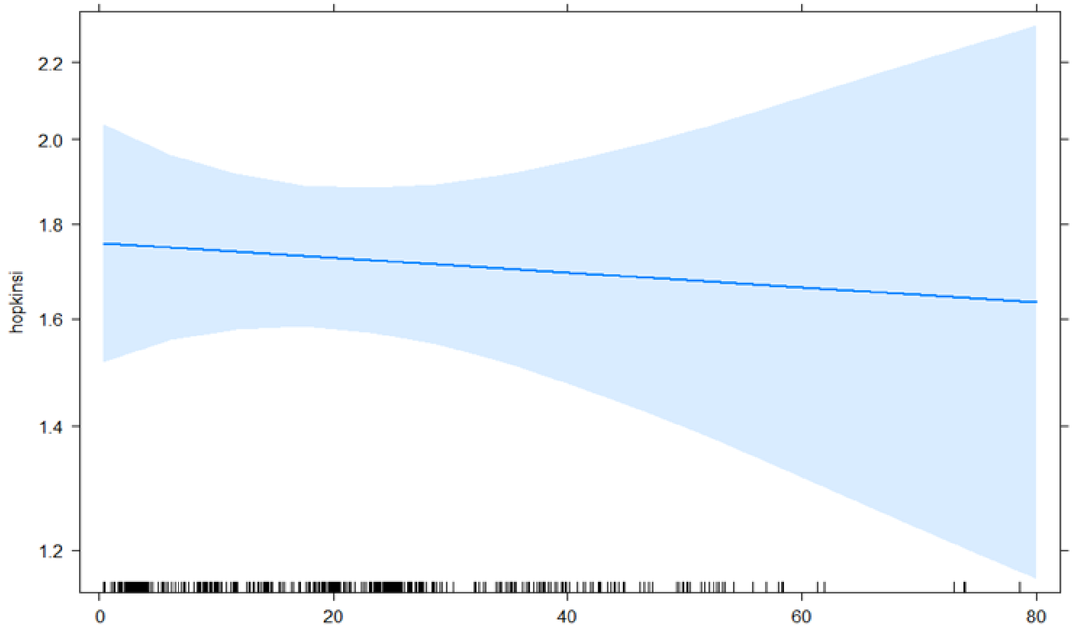

Figure 7. Effect of distance to reef edge on Squarespot abundance at Point Lobos.

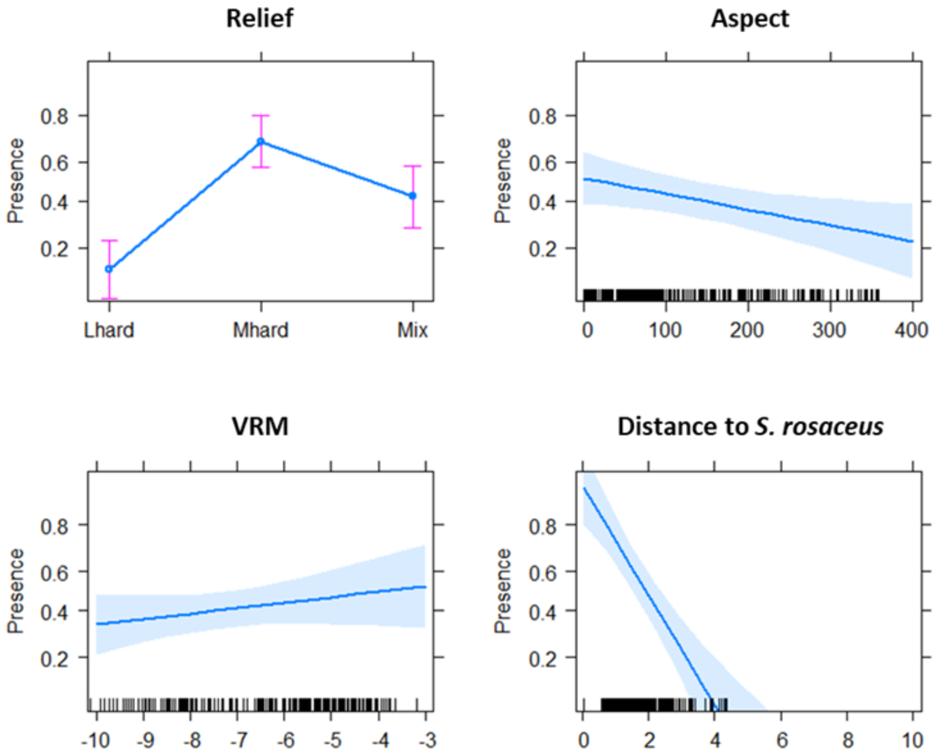

Figure 8. Effects plot of covariates predicting Squarespot occupancy at Point Sur. Note that the variable VRM has been log-transformed, and the variable distance to nearest Rosy rockfish has been cube-root transformed.

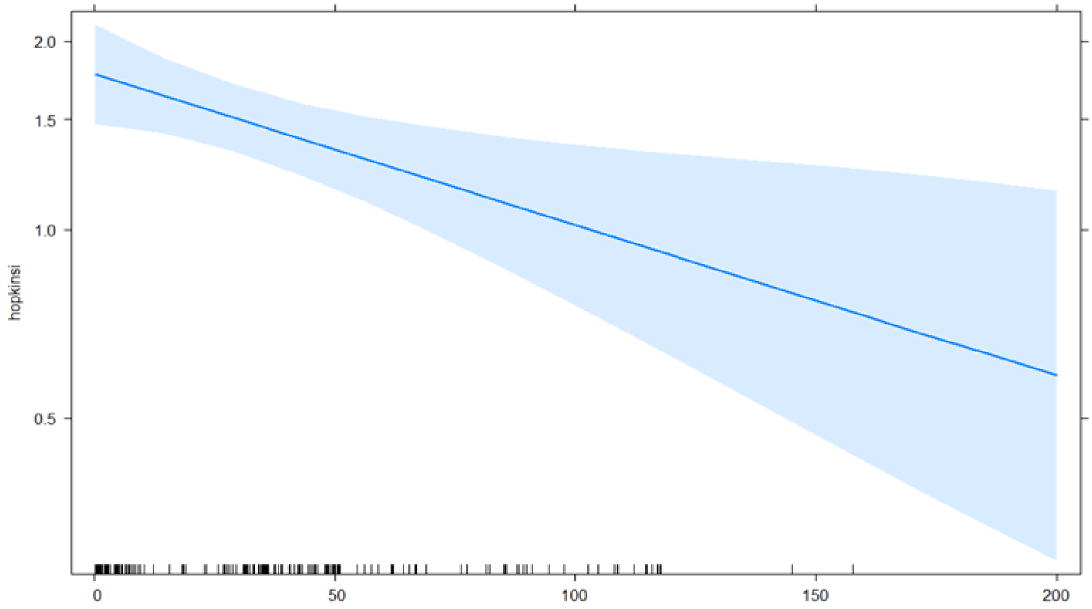

Figure 9. Model effects plot of distance to reef edge in predicting Squarespot abundance at Point Sur.

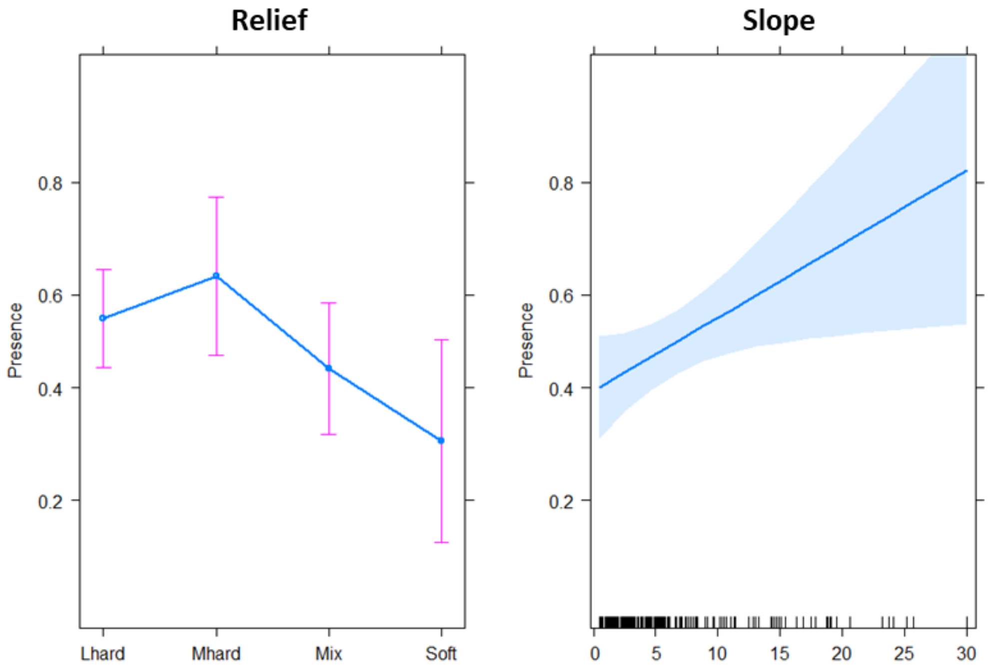

Figure 10. Model effects plot showing effects of variables on Rosy rockfish occupancy at Portuguese Ledge.

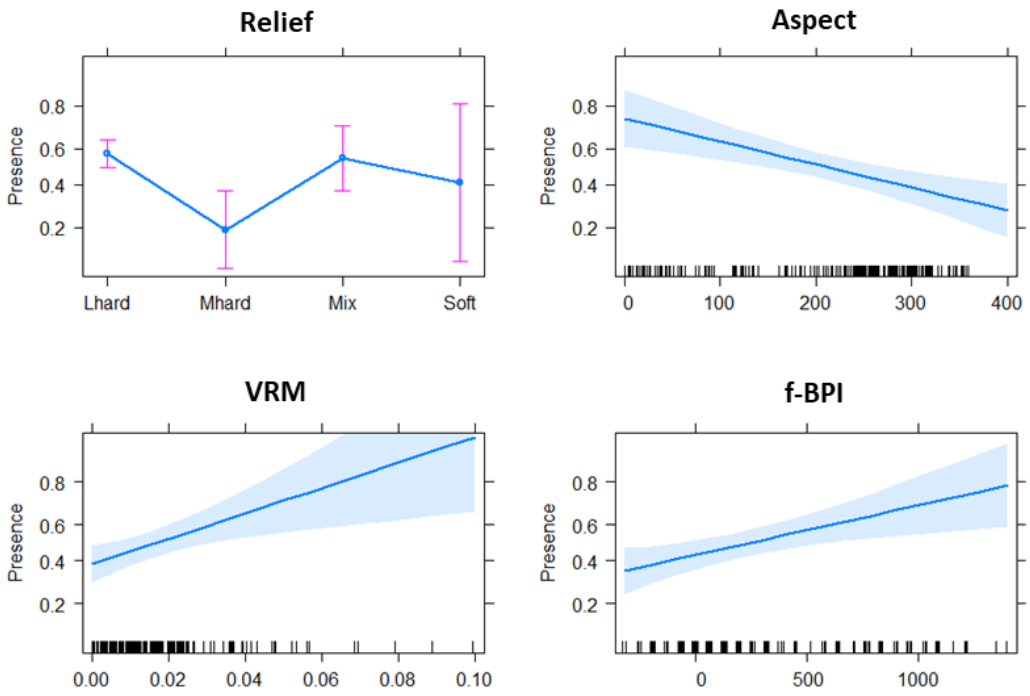

Figure 11. Effects of relief, aspect, VRM, and f-BPI on the probability of Rosy rockfish occupancy at Point Lobos. Note the habitat covariate VRM has been log-transformed.

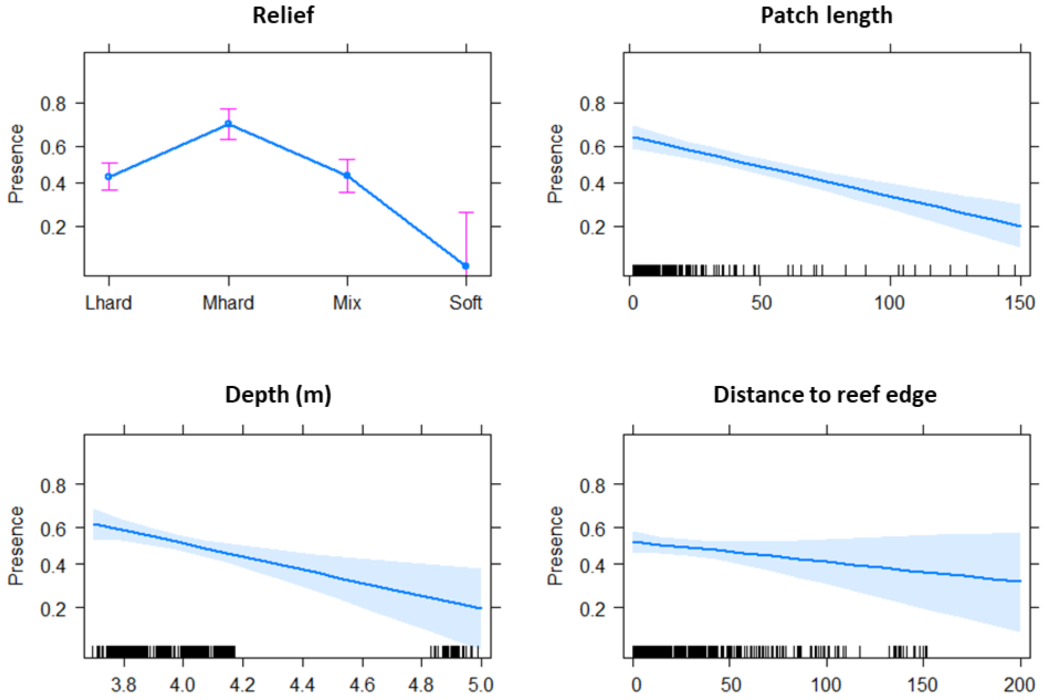

Figure 12. Effects of relief, patch length, depth, and distance to reef edge on the probability of Rosy rockfish presence at Point Sur. Note Depth has been log-transformed.

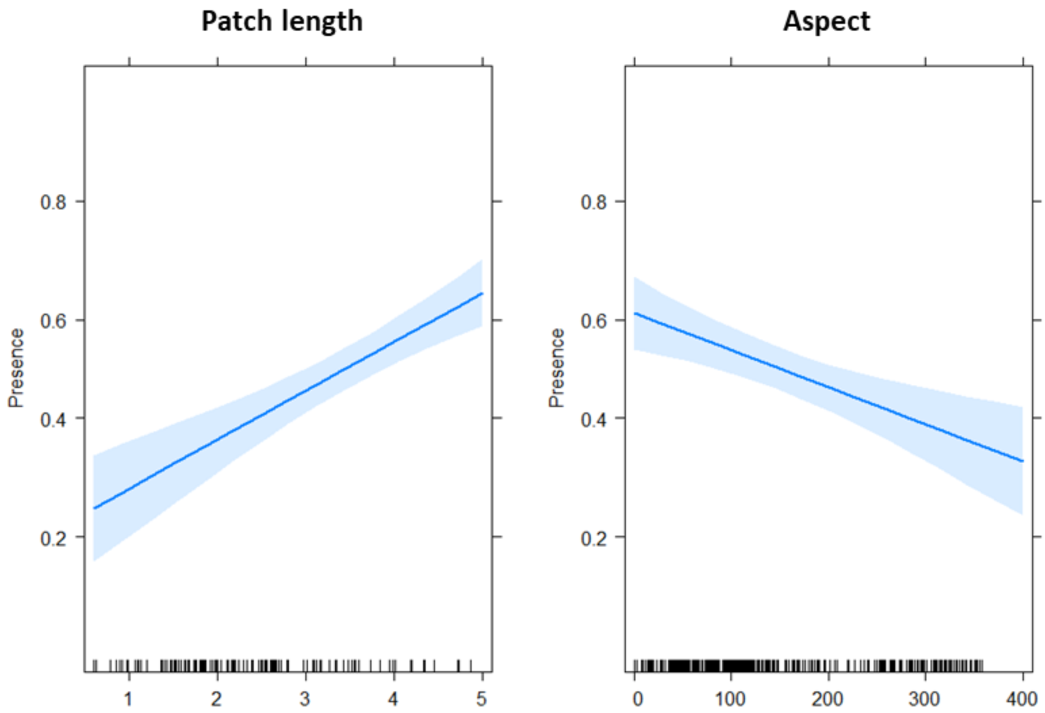

Figure 13. Effects of patch length and aspect on the probability of Pygmy rockfish occupancy at Portuguese Ledge. Note that the variable patch length has been log-transformed.

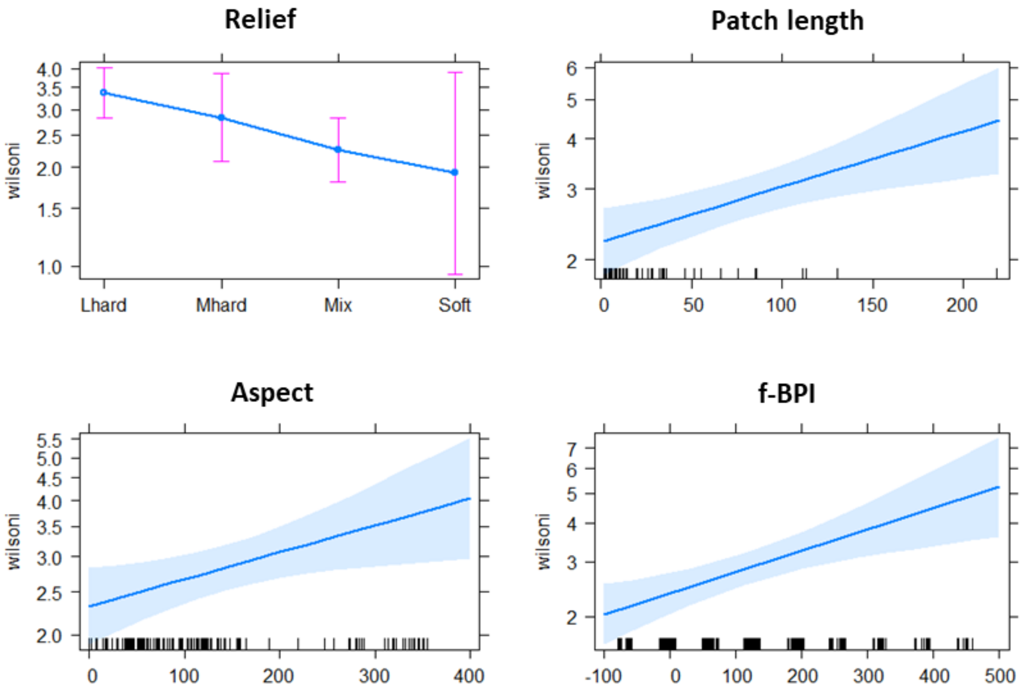

Figure 14. Effects of covariates on Pygmy abundance at Portuguese Ledge.

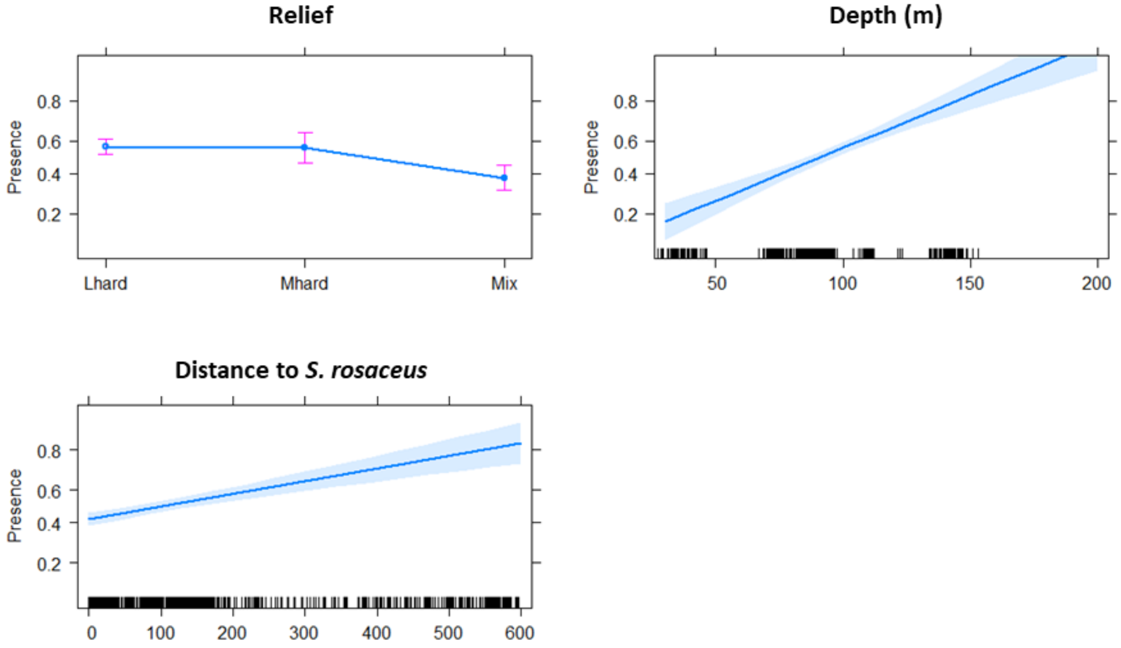

Figure 15. Effects of relief, depth, and distance to nearest Rosy rockfish on the probability of Pygmy presence at Point Lobos.

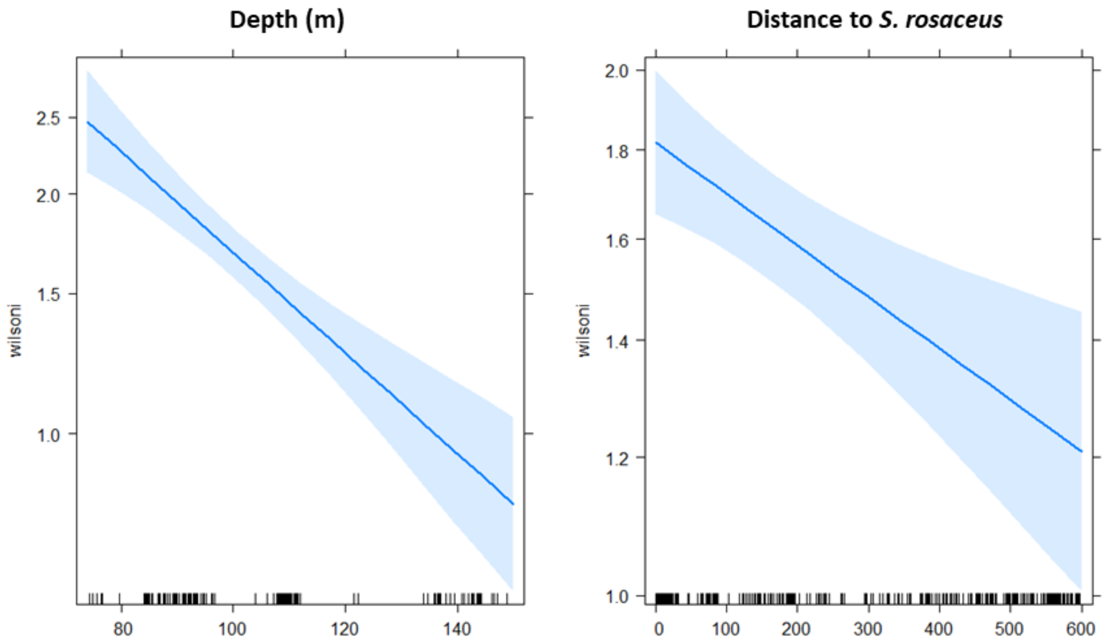

Figure 16. Effects of depth and distance to nearest Rosy rockfish on Pygmy abundance at Pt. Lobos.

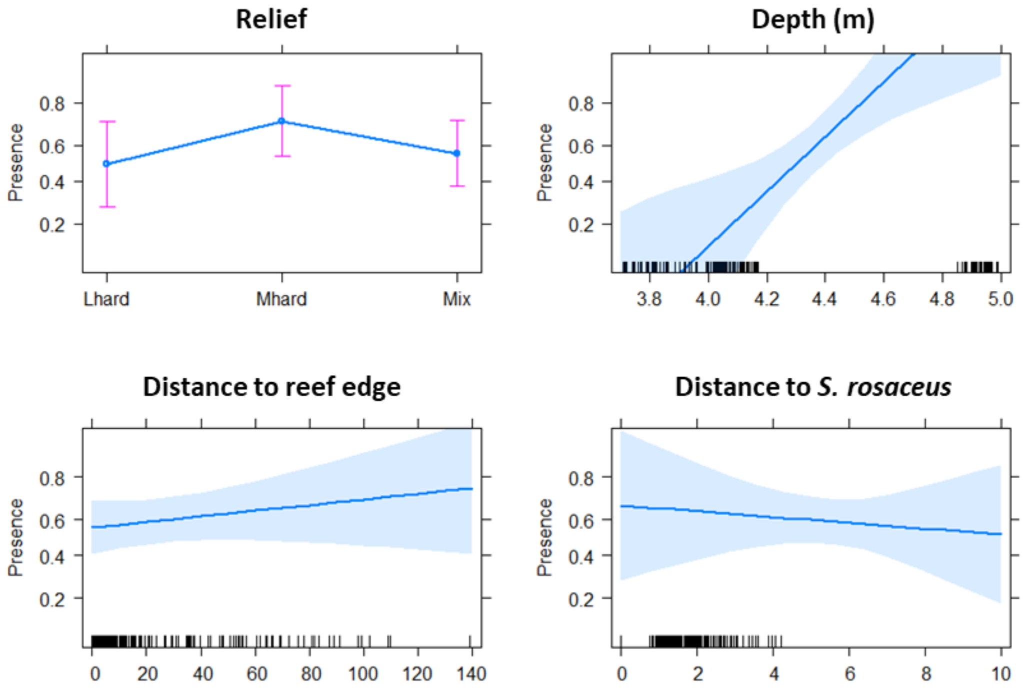

Figure 17. Model covariate effects on Pygmy occupancy at Pt. Sur. Note the habitat variable Depth has been log-transformed, and the variable Distance to Rosy rockfish has been cube-root transformed.

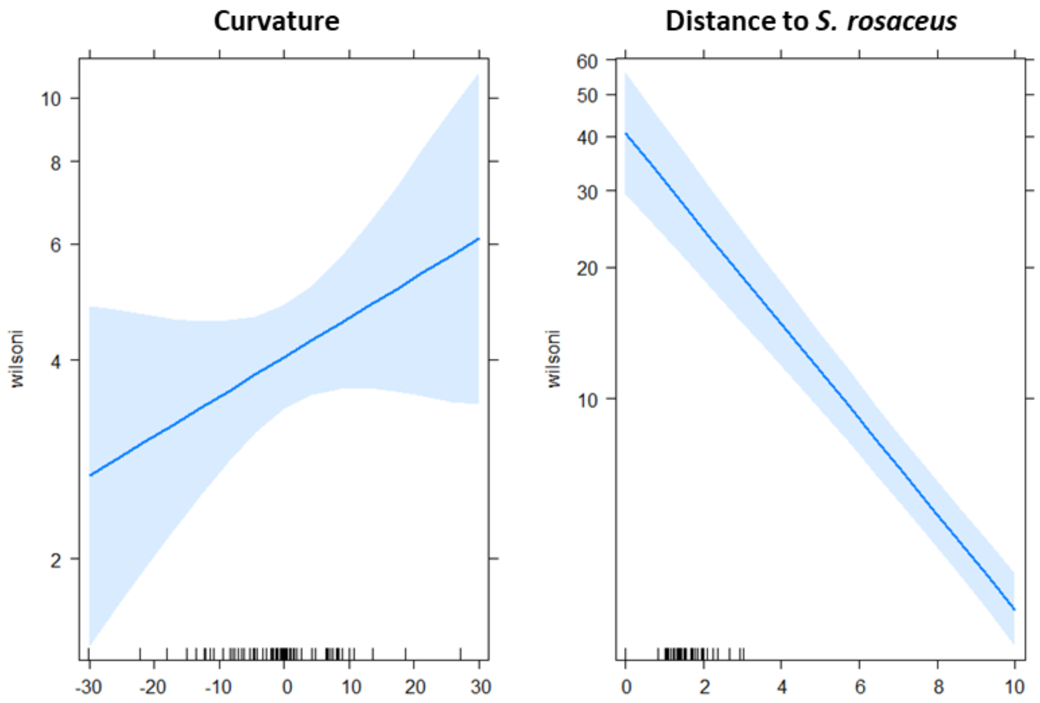

Figure 18. Model effects of curvature and distance to nearest Rosy rockfish on Pygmy rockfish abundance at Point Sur. Note distance to nearest Rosy rockfish has been cube-root transformed.

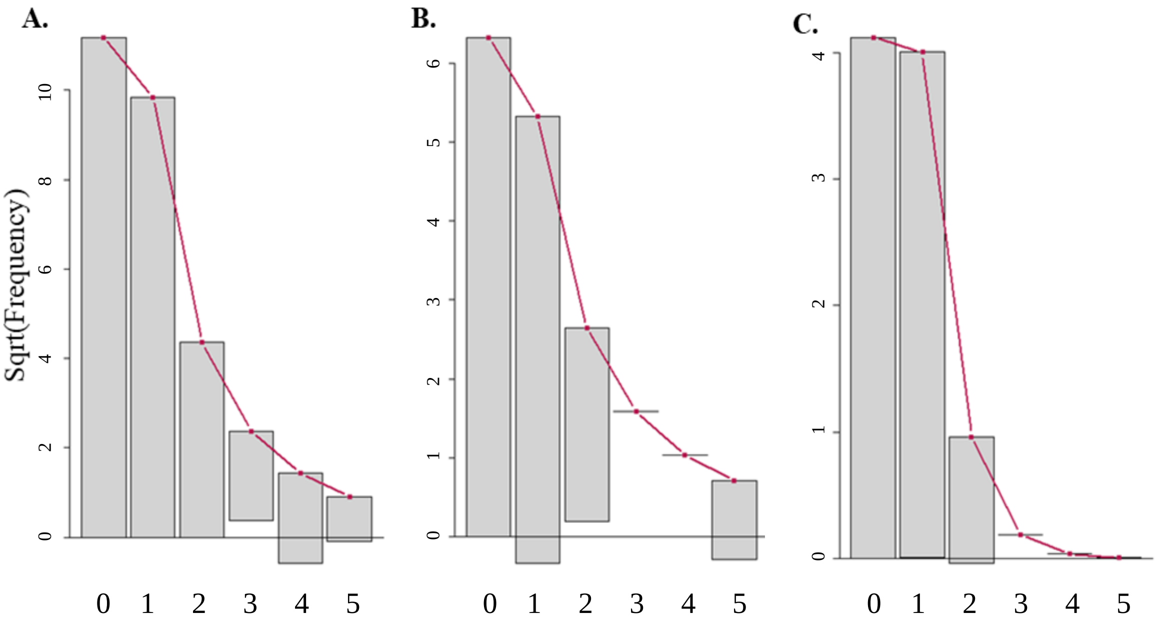

Figure 19. Rootograms of best model fit to Squarespot test data from (A) Portuguese Ledge, (B) Point Lobos, and (C) Point Sur. Frequencies have been square-root transformed to ensure infrequent count categories are not overlooked.

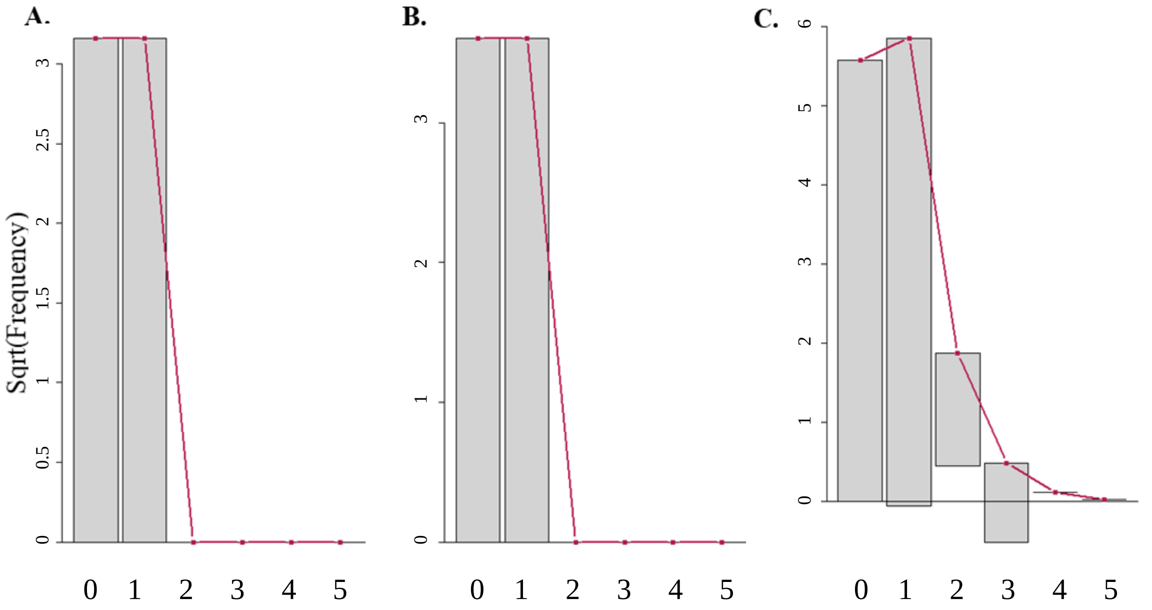

Figure 20. Rootograms of best model fit to Rosy rockfish test data from (A) Portuguese Ledge, (B) Point Lobos, and (C) Point Sur.

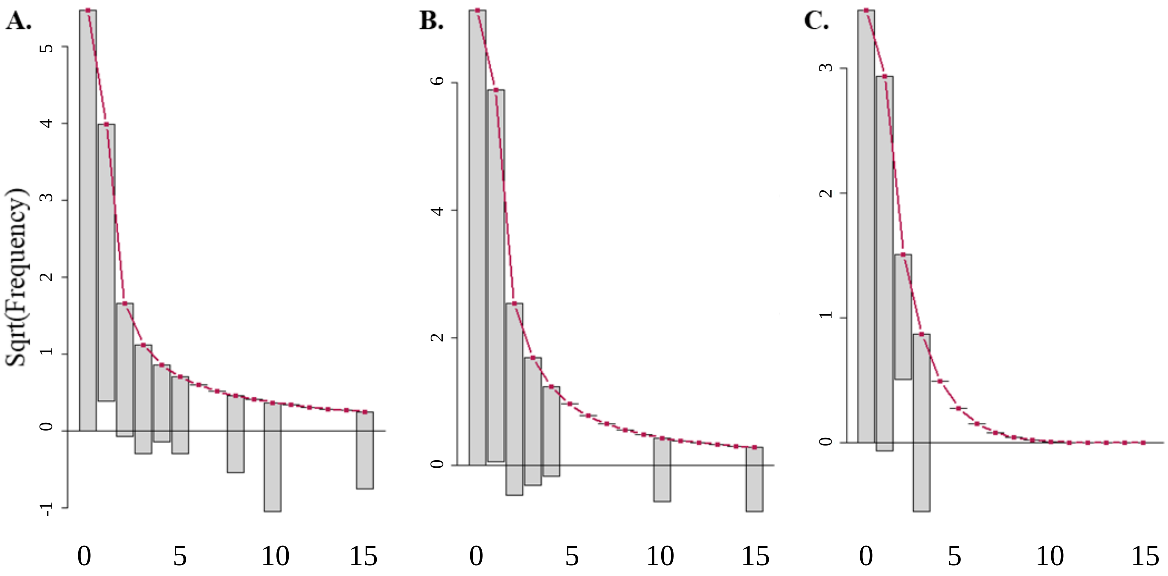

Figure 21. Rootograms of best model fit to Pygmy test data from (A) Portuguese Ledge, (B) Point Lobos, and (C) Point Sur.

REFERENCES

Anscombe, F. J., & Tukey, J. W. (1963). The examination and analysis of residuals. *Technometrics*, *5*(2), 141-160.

Anderson, T. J., & Yoklavich, M. M. (2007). Multiscale habitat associations of deepwater demersal fishes off central California. *Fishery Bulletin*, *105*(2), 168-179.

Anderson, T. J., Syms, C., Roberts, D. A., & Howard, D. F. (2009). Multi-scale fish–habitat associations and the use of habitat surrogates to predict the organisation and abundance of deep-water fish assemblages. *Journal of Experimental Marine Biology and Ecology*, *379*(1-2), 34-42.

Auster, P. J., Malatesta, R. J., & LaRosa, S. C. (1995). Patterns of microhabitat utilization by mobile megafauna on the southern New England (USA) continental shelf and slope. *Marine Ecology Progress Series*, *127*, 77-85.

Barss, W., Hixon, M. A., Stein, D. L., & Tissot, B. N. (1992). Fish-habitat associations on a deep reef at the edge of the Oregon continental shelf.

Barnston, A. G. (1992). Correspondence among the correlation, RMSE, and Heidke forecast verification measures; refinement of the Heidke score. *Weather and Forecasting*, *7*(4), 699-709.

Beck, M.W. (2000). Separating the elements of habitat structure: independent effects of habitat complexity and structural components on rocky intertidal gastropods. Journal of Experimental Marine Biology and Ecology, 249(1):29-49.

Bivand, R. S., & Wong, D. W. (2018). Comparing implementations of global and local indicators of spatial association. *Test*, *27*(3), 716-748.

Boucher, J. P., Denuit, M., & Guillén, M. (2007). Risk classification for claim counts: a comparative analysis of various zeroinflated mixed Poisson and hurdle models. *North American Actuarial Journal*, *11*(4), 110-131.

Carlevaro, F., Croissant, Y., & Hoareau, S. (2012). Multiple hurdle models in R: The mhurdle Package. *URL: http://cran. r-project. org/web/packages/mhurdle/vignettes/mhurdle. pdf*.

Chai, T., & Draxler, R. R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)?–Arguments against avoiding RMSE in the literature. *Geoscientific model development*, *7*(3), 1247-1250.

Eschmeyer, W.N., E.S. Herald and H. Hammann, 1983. A field guide to Pacific coast fishes of North America. Boston (MA, USA): Houghton Mifflin Company. xii+336 p.

ESRI 2011. ArcGIS Desktop: Release 10. Redlands, CA: Environmental Systems Research Institute.

Haldorson, L., & Love, M. (1991). Maturity and fecundity in the rockfishes, Sebastes spp., a review. *Marine Fisheries Review*, *53*(2), 25-31.

Harvey, C. J., Tolimieri, N., & Levin, P. S. (2006). Changes in body size, abundance, and energy
allocation in rockfish assemblages of the northeast Pacific. Ecological Applications, 16(4), 1502-1515.

Huff, D. D., Yoklavich, M. M., Love, M. S., Watters, D. L., Chai, F., & Lindley, S. T. (2013). Environmental factors that influence the distribution, size, and biotic relationships of the Christmas tree coral Antipathes dendrochristos in the Southern California Bight. *Marine Ecology Progress Series*, *494*, 159-177.

Iampietro, P. J., Young, M. A., & Kvitek, R. G. (2008). Multivariate prediction of rockfish habitat suitability in Cordell bank national marine sanctuary and Del Monte Shalebeds, California, USA. *Marine Geodesy*, *31*(4), 359-371.

Ierodiaconou, D., Monk, J., Rattray, A., Laurenson, L., & Versace, V. L. (2011). Comparison of automated classification techniques for predicting benthic biological communities using hydroacoustics and video observations. *Continental Shelf Research*, *31*(2), S28-S38.

Jackman, S. (2010). pscl: Classes and methods for R. Developed in the Political Science Computational Laboratory, Stanford University. Department of Political Science, Stanford University, Stanford, CA. R package version 1.03. 5. *http://www. pscl. stanford. edu/.*

Johnson, S. W., Murphy, M. L., & Csepp, D. J. (2003). Distribution, habitat, and behavior of
rockfish, *Sebastes* spp., in nearshore waters of southeastern Alaska: observations from a remotely operated vehicle. Environmental Biology of Fishes, 66, 259-270.

Karpov, K. A., Albin, D. P., & Van Buskirk, W. H. (1995). *The marine recreational fishery in northern and central California: A historical comparison (1958-86), status of stocks (1980-86), and effects of changes in the California Current* (No. 176). State of California, Resources Agency, Department of Fish and Game.

Kostylev, V. E., Erlandsson, J., Ming, M. Y., & Williams, G. A. (2005). The relative importance of habitat complexity and surface area in assessing biodiversity: fractal application on rocky shores. *Ecological Complexity*, *2*(3), 272-286.

Laidig, T. E., Watters, D. L., & Yoklavich, M. M. (2009). Demersal fish and habitat associations from visual surveys on the central California shelf. *Estuarine, coastal and shelf science*, *83*(4), 629-637.

Legendre, P. (1993). Spatial autocorrelation: trouble or new paradigm?. *Ecology*, *74*(6), 1659-1673.

Levin, P. S., Holmes, E. E., Piner, K. R., & Harvey, C. J. (2006). Shifts in a Pacific Ocean fish
assemblage: the potential influence of exploitation. Conservation Biology, 20(4), 1181-1190.

Love, M. S. (1990). Life history aspects of 19 rockfish species (Scorpaenidae: Sebastes) from the Southern California Bight.

Love, M. S., Yoklavich, M., Thorsteinson, L., & Buttler, J. (2002). *The rockfishes of the Northeast Pacific*. Berkeley: University of California Press.

Love, M. S., & Yoklavich, M. (2006). Deep rock habitats. *The ecology of marine fishes: California and adjacent waters*, 253-266.

Love, M. S., & York, A. (2006). The relationships between fish assemblages and the amount of bottom horizontal beam exposed at California oil platforms: fish habitat preferences at man-made platforms and (by inference) at natural reefs. *Fishery Bulletin*, *104*(4), 542-550.

Love, M. S., Yoklavich, M., & Schroeder, D. M. (2009). Demersal fish assemblages in the Southern California Bight based on visual surveys in deep water. *Environmental Biology of Fishes*, *84*(1), 55-68.

Lundblad, E. M. I. L. Y., Wright, D. J., Miller, J. O. Y. C. E., Larkin, E. M., Rinehart, R. O. N. A. L. D., Battista, T., ... & Donahue, B. T. (2006). Classifying benthic terrains with multibeam bathymetry, bathymetric position and rugosity: Tutuila, American Samoa. *Mar. Geod*, *29*(2), 89-111.

Martin, J., & Hall, D. B. (2016). R 2 measures for zero-inflated regression models for count data with excess zeros. *Journal of Statistical Computation and Simulation*, *86*(18), 3777-3790.

Mason, J. E. (1998). Declining rockfish lengths in the Monterey Bay, California, recreational fishery, 1959–94. *Marine Fisheries Review*, *60*(3), 15-28.

McClatchie, S., Millar, R. B., Webster, F., Lester, P. J., Hurst, R., & Bagley, N. (1997). Demersal fish community diversity off New Zealand: is it related to depth, latitude and regional surface phytoplankton?. *Deep Sea Research Part I: Oceanographic Research Papers*, *44*(4), 647-667.

Miller, D. J., and J. J. Geibel. 1973. Summary of blue rockfish and lingcod life histories; a reef ecology study; and giant kelp, *Macrocystis pyrifera*, experiments in Monterey Bay, California. Calif. Dep. Fish Game, Fish Bull. 158, 137 p.

Montgomery DC, Peck EA (1992) Introduction to linear regression analysis, 2nd edn, Probability and Statistics Series. John Wiley & Sons, New York, NY

Morgan, L. E., Etnoyer, P., Scholz, A. J., Mertens, M., & Powell, M. (2005). Conservation and management implications of deep-sea coral and fishing effort distributions in the Northeast Pacific Ocean. In A. Friewald, & J. M. Roberts (Eds.), Cold-water Corals and Ecosystems (pp. 1171-1187). Springer-Verlag Berlin Heidelberg.

Morrison, R. E., Bryant, C. M., Terejanu, G., Prudhomme, S., & Miki, K. (2013). Data partition methodology for validation of predictive models. *Computers & Mathematics with Applications*, *66*(10), 2114-2125.

National Atmospheric & Oceanographic Administration. (2010). NOAA strategic plan for deep-sea coral and sponge ecosystems: research, management, and international cooperation. NOAA Coral Reef and Conservation Program. NOAA Technical Memorandum CRCP 11, 67.

[OCNMS] Olympic Coast National Marine Sanctuary. (2011). Final Management Plan and Environmental Assessment. 329 p.

Ord, J. K., & Getis, A. (1995). Local spatial autocorrelation statistics: distributional issues and an application. *Geographical analysis*, *27*(4), 286-306.

Pacunski, R. E., & Palsson, W. A. (2001). Macro- and micro-habitat relationships of sub-adult and adult rockfish, lingcod, and kelp greenling in Puget Sound. Puget Sound Research 2001 Proceedings. Puget Sound Water Quality Action Team, pp. 1-11.

Parker, S. J., Berkeley, S. A., Golden, J. T., Gunderson, D. R., Heifetz, J., Hixon, M. A., ... & O'Connell, V. M. (2000). Management of Pacific rockfish. *Fisheries*, *25*(3), 22-30.

Pinheiro, J. C., & Bates, D. M. (2000). Mixed-effects models in S and S-Plus.,(Springer-Verlag: New York). *Mixed-effects models in S and S-Plus. Springer Verlag, New York.*

Posada D, Buckley TR (2004) Model selection and model averaging in phylogenetics: advantages of Akaike Information Criterion and Bayesian approaches over Likelihood Ratio Tests. Syst Biol 53:793–808

R Development Core Team (2010) R: A language and environment for statistical computing. R Foundation for Statisitcal Computing, Vienna. http://www.R-project.org/.

Rotenberry, J. T., Preston, K. L., & Knick, S. T. (2006). GIS-BASED NICHE MODELING FOR MAPPING SPECIES' HABITAT. *Ecology*, *87*(6), 1458-1464.

Sakamoto, Y., Ishiguro, M., & Kitagawa, G. (1986). Akaike information criterion statistics. *Dordrecht, The Netherlands: D. Reidel*, *81*.

Segurado, P. A. G. E., Araujo, M. B., & Kunin, W. E. (2006). Consequences of spatial autocorrelation for niche-based models. *Journal of Applied Ecology*, *43*(3), 433-444.

Starr, R. M., Carr, M. H., Caselle, J., Estes, J. A., Pomeroy, C., Syms, C., Yoklavich, M. M.
(2002). A review of the ecological effectiveness of subtidal marine reserves in central California Part I: synopsis of scientific investigations. Monterey Bay National Marine Sanctuary, 1-131.

Stein, D. L. (1992). Fish-habitat associations on a deep reef at the edge of the Oregon continental shelf. *Fish. Bull.*, *90*, 540-551.

Tissot, B. N., Yoklavich, M. M., Love, M. S., York, K., & Amend, M. (2006). Benthic invertebrates that form habitat on deep banks off southern California, with special reference to deep sea coral. *Fishery Bulletin*, *104*(2), 167-181.

Valentine, P.C., S.J. Fuller, and L.A. Scully. (2004). Terrain Ruggedness Analysis and Distribution of Boulder Ridges in the Stellwagen Bank National Marine Sanctuary Region (poster). Galway, Ireland: 5th International Symposium on Marine Geological and Biological Habitat Mapping (GeoHAB), May 2004.

Weiss, A. D. (2001). Topographic Positions and Landforms Analysis (poster), ESRI International User Conference, July 2001. San Diego, CA: ESRI.

Whitmire, C. E., Clarke, M. E., Yoklavich, M. M., Everett, M. V., Hourigan, T. F., & Cairns, S. D. (2017). Deep-Sea Coral Taxa in the US West Coast Region: Depth and Geographical Distribution.

Willmott, C. J., & Matsuura, K. (2005). Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Climate research*, *30*(1), 79-82.

Wilson, M.F.J., B. O'Connell, C. Brown, J.C. Guinan, and A.J. Grehan. (2007). Multiscale terrain analysis of multibeam bathymetry data for habitat mapping on the continental slope. Marine Geodesy 30, 3–35.

Yoklavich, M. M., Greene, H. G., Cailliet, G. M., Sullivan, D. E., Lea, R. N., & Love, M. S. (2000). Habitat associations of deep-water rockfishes in a submarine canyon: an example of a natural refuge. *Fishery Bulletin*, *98*(3), 625-641.

Yoklavich, M. M., Love, M. S., & Forney, K. A. (2007). A fishery-independent assessment of an overfished rockfish stock, cowcod (Sebastes levis), using direct observations from an occupied submersible. *Canadian Journal of Fisheries and Aquatic Sciences*, *64*(12), 1795-1804.

Zeileis, A., Kleiber, C., & Jackman, S. (2008). Regression models for count data in R. *Journal of statistical software*, *27*(8), 1-25.

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

Zuur, A. F., Ieno, E. N., & Elphick, C. S. (2010). A protocol for data exploration to avoid common statistical problems. *Methods in ecology and evolution*, *1*(1), 3-14.

APPENDICES

Appendix A. Model output summarizing coefficients and significance (p-values) of each habitat variable used in the model predicting Squarespot occurrence at Portuguese Ledge. Significant codes: * = < 0.05, ** = <0.005, *** = <0.001.

Parameter	coefficient	Std. Error	z-value	p-value
Intercept	18.08	2.898	6.239	<0.001***
Mhard	1.090	0.4972	2.193	0.02831*
Mix	0.4919	0.3432	1.433	0.1518
Soft	-0.2068	0.5977	-0.346	0.7294
Aspect	-0.0041	0.0014600	-2.836	0.004570**
Depth	-0.1987	0.03224	-6.163	<0.001***

Appendix B. Model output summarizing coefficients and significance (p-values) of each habitat variable used in the model predicting Squarespot abundance at Portuguese Ledge.

Parameter	coefficient	Std. Error	z-value	p-value
Intercept	4.322	24.44	0.1770	0.8596
depth	-0.1081	0.04511	-2.396	0.01660*

Appendix C. Model output summarizing coefficients and significance (p-values) of each habitat variable used in the model predicting Squarespot occurrence at Point Lobos.

Parameter	coefficient	Std. Error	z-value	p-value
Intercept	0.5392	0.0989	5.452	<0.001***
Mhard	0.0213	0.1910	0.1065	0.9152
Mix	-1.431	0.2238	-6.393	<0.001***
r_dist	-0.002641	0.0006042	-4.3709	<0.001***

Appendix D. Model output summarizing coefficients and significance (p-values) of each habitat variable used in the model predicting Squarespot abundance at Point Lobos.

Parameter	coefficient	Std. Error	z-value	p-value
Intercept	-11.46	65.12	-0.1760	0.8604
edge_dist	-0.003834	0.008495	-0.4512	0.6518

Appendix E. Model output summarizing coefficients and significance (p-values) of each habitat variable used in the model predicting Squarespot occurrence at Point Sur.

Parameter	coefficient	Std. Error	z-value	p-value
Intercept	2.025	0.7184	2.818	0.004832**
Mhard	2.266	0.4183	6.349	<0.001***
Mix	1.510	0.3967	3.806	<0.001***
aspect	-0.003287	0.001359	-2.419	0.01556*
log_VRM	0.1137	0.09348	1.216	0.2240
cube_r_dist	-0.1082	0.02577	-4.199	<0.001***

Appendix F. Model output summarizing coefficients and significance (p-values) of each habitat variable used in the model predicting Squarespot abundance at Point Sur.

Parameter	coefficient	Std. Error	z-value	p-value
Intercept	-10.51	79.38	-0.1324	0.8947
edge_dist	-0.0202	0.005694	-3.553	<0.001***

Appendix G. Model output summarizing coefficients and significance (p-values) of each habitat variable used in the model predicting Rosy occurrence at Portuguese Ledge.

Parameter	coefficient	Std. Error	z-value	p-value
Intercept	-0.3255	0.3552	-0.9163	0.3595
Mhard	0.4073	0.4468	0.9116	0.3620
Mix	-0.2389	0.3969	-0.6019	0.5472
Soft	-0.9498	0.5540	-1.745	0.08644
slope	0.0674	0.0297	2.269	0.0233*

Appendix H. Model output summarizing coefficients and significance (p-values) of each habitat variable used in the model predicting Rosy abundance at Portuguese Ledge.

Parameter	coefficient	Std. Error	z-value	p-value
Intercept	-2.823	0.5745	-4.914	<0.001***

Appendix I. Model output summarizing coefficients and significance (p-values) of each habitat variable used in the model predicting Rosy occurrence at Point Lobos.

Parameter	coefficient	Std. Error	z-value	p-value
Intercept	0.7323	0.3819	1.917	0.0551
Mhard	-1.837	0.5356	-3.430	<0.001***
Mix	-0.455	0.4252	-1.071	0.2843
Soft	-0.8665	1.180	-0.7340	0.4629
aspect	-0.004780	0.001480	-3.233	0.001224**
VRM	19.95	9.252	2.156	0.03107*
f-BPI	0.0009700	0.0004200	2.345	0.01901*

Appendix J. Model output summarizing coefficients and significance (p-values) of each habitat variable used in the model predicting Rosy abundance at Point Lobos.

Parameter	coefficient	Std. Error	z-value	p-value
Intercept	-3.407	0.7052	-4.831	<0.001***

Appendix K. Model output summarizing coefficients and significance (p-values) of each habitat variable used in the model predicting Rosy occurrence at Point Sur.

Parameter	coefficient	Std. Error	z-value	p-value
Intercept	1.785	0.4071	4.385	<0.001***
patch length	-0.01209	0.002250	-5.377	<0.001***
Mhard	1.317	0.2582	5.101	<0.001***
Mix	0.001580	0.2703	0.006000	0.9953
Soft	-1.756	0.6642	-2.643	0.008211**
depth	-0.02377	0.006840	-3.475	<0.001***
edge_dist	-0.006590	0.003390	-1.942	0.05216

Appendix L. Model output summarizing coefficients and significance (p-values) of each habitat variable used in the model predicting Rosy abundance at Point Sur.

Parameter	coefficient	Std. Error	z-value	p-value
Intercept	-1.679	0.1799	-9.335	<0.001***

Appendix M. Model output summarizing coefficients and significance (p-values) of each habitat variable used in the model predicting Pygmy occurrence at Portuguese Ledge.

Parameter	coefficient	Std. Error	z-value	p-value
Intercept	-0.1313	0.1628	-0.8060	0.4200
patch length	0.008690	0.001650	5.284	<0.001***
aspect	-0.002280	0.0009200	-2.480	0.03100*

Appendix N. Model output summarizing coefficients and significance (p-values) of each habitat variable used in the model predicting Pygmy abundance at Portuguese Ledge.

Parameter	coefficient	Std. Error	z-value	p-value
Intercept	-1.434	0.6003	-2.389	0.01689*
Mhard	0.3688	0.3471	1.063	0.2879
Mix	-0.5349	0.3005	-1.780	0.07500
Soft	-0.7418	0.6373	-1.164	0.2445
patch length	0.003730	0.001840	2.024	0.04293*
aspect	0.004130	0.001080	3.818	<0.001***
f-BPI	0.001730	0.008750	1.974	0.04838*

Appendix O. Model output summarizing coefficients and significance (p-values) of each habitat variable used in the model predicting Pygmy occurrence at Point Lobos.

Parameter	coefficient	Std. Error	z-value	p-value
Intercept	-2.573	0.3052	-8.429	<0.001***
depth	0.02578	0.003550	7.260	<0.001***
Mhard	-0.06475	0.2364	-0.2740	0.7840
Mix	-0.8099	0.2062	-3.928	<0.001***
r_dist	0.003280	0.0005100	6.458	<0.001***

Appendix P. Model output summarizing coefficients and significance (p-values) of each habitat variable used in the model predicting Pygmy abundance at Point Lobos.

Parameter	coefficient	Std. Error	z-value	p-value
Intercept	3.246	0.5061	6.415	<0.001***
depth	-0.03195	0.005420	-5.894	<0.001***
r_dist	-0.001860	0.0004700	-3.966	<0.001***

Appendix Q. Model output summarizing coefficients and significance (p-values) of each habitat variable used in the model predicting Pygmy occurrence at Point Sur.

Parameter	coefficient	Std. Error	z-value	p-value
Intercept	-8.176	1.945	-4.205	<0.001***
Mhard	-0.3705	0.5491	-0.6750	0.4999
Mix	-0.9916	0.7284	-1.361	0.1734
depth	0.1291	0.03617	3.568	<0.001***
edge_dist	0.005520	0.007320	0.7550	0.4502
r_dist	-0.1547	0.07148	-2.165	0.03040*

Appendix R. Model output summarizing coefficients and significance (p-values) of each habitat variable used in the model predicting Pygmy abundance at Point Sur.

Parameter	coefficient	Std. Error	z-value	p-value
Intercept	1.397	0.1361	10.27	<0.001***
curvature	0.01709	0.009460	1.807	0.07080
r_dist	-0.03301	0.004800	-6.874	<0.001***