Modeling Rockfish Abundance and Distribution on Cordell Bank National Marine Sanctuary, California using Generalized Linear Models (GLMs)

A Capstone Project

Presented to the Faculty of Science and Environmental Policy

in the

College of Science, Media Arts, and Technology

at

California State University, Monterey Bay

in Partial Fulfillment of the Requirements for the Degree of

Bachelor of Science

by

Mary Young

May 1, 2007

TABLE OF CONTENTS

Page

ACKNOWLEDGEMENTS

ABSTRACT

Due to the decline of fisheries throughout the world, there is an increasing demand for more stock assessment data by fisheries managers. Along the Pacific coast of North America there are insufficient stock data for most rockfish species, which compose one of the most valuable commercial and recreational fisheries in California. One approach being explored for more efficiently generating stock assessment data over large areas is the use of habitat-based community assessment. The general hypothesis is that because rockfish are not randomly distributed across habitats, it should be possible to model and predict their distribution and abundance based on habitat maps and biological data. The purpose of this study is to test this hypothesis using autoclassification of multibeam bathymetry and acoustic backscatter data along with submersible video data of the seafloor and rockfish from Cordell Bank National Marine Sanctuary to determine how well rockfish distribution of three species of rockfish (*Sebastes flavidus* (yellowtail), *S. rosaceus* (rosy), and *S. elongatus* (greenstriped)) can be modeled based on seafloor habitat parameters. In addition, those methods were compared to methods using the video data alone. General linear models (GLMs) were created using rugosity, slope, aspect, depth, and topographic position index analyses of bathymetric digital elevation models and supervised surface texture classification from the backscatter mosaic along with the presence/absence points for the three species of rockfish. These models proved to be the most efficient at accurately predicting the distribution of *S. rosaceus* with an average accuracy of 81%. The GLMs correctly predicted the occurrence of *S. flavidus* 76% of the time and the occurrence of *S. elongatus* 62% of the time. When compared to the analysis of the video results, the GLMs were better at predicting the abundance of *S. flavidus* and *S. rosaceus* with percent errors of 16% and 13% respectively. Both methods were equal for *S. elongatus*. The techniques used in this study could be used for management purposes such as planning and evaluating locations for marine protected areas.

LIST OF TABLES

LIST OF FIGURES

1 INTRODUCTION

Fisheries throughout the world are in decline. Currently, all of the world's seventeen oceanic fisheries are at or below sustainable yield (Wilson, 2002). In addition, the Food and Agricultural Organization of the United Nations (FAO) reports that "69% of the world's [fish] stocks…are either fully to heavily exploited, overexploited, or depleted…and therefore are in need of urgent conservation and management measures" (Lauck et al., 1998). This is mostly due in part to the human mismanagement of this resource (Clark, 1996).

Along the Pacific Coast of the United States, rockfish make up one of the most economically important fisheries in California. However, because they are slow-growing, long-lived, and old at reproductive maturity, they are extremely vulnerable to fishing pressure. Recent findings show that some of these species are being harvested at unsustainable levels. Although there is a need, it is difficult to manage these fisheries because there are insufficient data on the distribution, abundance, and habitat characteristics of the different species (Yoklavich et al., 1999).

Therefore, in order to create better management policies, there is a call for more and better stock assessment data. Along the Pacific Coast, only 22% of the rockfish species have ever been assessed and many of them are commercially important species. In addition, from the rockfish that were assessed, half of them were reported as "overfished" or "approaching overfished condition". This lack of data is mostly due in part to the inefficiency of single species stock assessments (Nasby-Lucas et al., 2002). There are limitations in the ability of many of the field sampling techniques to get an accurate stock representation. For example, bottom trawls are extensively used to determine the abundance of fish along the continental shelf and some offshore rocky banks. However, this technique precludes enumeration of many of the fish on top of banks and has proven difficult to use in any rocky area (Starr et al., 1995).

As a possible solution to the problem of incomplete stock data, there is the growing use of habitat-based community assessment. This technique recognizes that species are not randomly distributed throughout their environment. Rather, the distribution and abundance of many rockfish species are associated with the assortment, quality, and extent of benthic habitats (Starr et al., 1995; Nasby-Lucas et al., 2002). Furthermore, in the marine environment, it has

been found that habitat properties are better at capturing the community-level and ecosystem-level schemes rather than relying on species richness and endemism (Leslie et al., 2003).

Species-specific habitat modeling based on remotely sensed seafloor geomorphology and texture data may prove to be an efficient technique for the stock assessment of rockfish and other groundfish species. This technique can help to distinguish between the different types of habitats as well as give information on the abundance and distribution of different species. With the advent of new technologies, sonar systems are now in existence that provide high enough resolution bathymetric and backscatter data for the classification of habitats over large areas (Nasby-Lucas et al., 2002). These data can be included with groundtruth data and fish observation data to provide an accurate habitat model that includes species abundance and distribution.

This approach has been applied with some success by a number of researchers in the terrestrial realm. Many studies have been completed in order to relate vegetation distribution and abundance to environmental parameters such as topographic position index, climate, aspect, rainfall patterns, etc. For example, Austin et al. (1994) used presence/absence data from nine different eucalyptus species in Australia and variables such as variability in rainfall, variability in temperature, seasonal levels of radiation, topographic position, lithology class, and nutrient index to create a statistical model to predict the probability of occurrence of this eucalyptus where presence/absence data was not available. The authors completed this analysis by using generalized linear models (GLMs). Using these models, they were able to accurately predict at least 90% of the presence/absence for 7 of the species and more than 80% of the presence/absence for the other two species. This study concluded that the equations from the GLMs provide adequate models for these nine species of eucalyptus (Austin et al., 1994).

In addition to terrestrial vegetation, these same modeling techniques have been used to relate mobile terrestrial animals to habitat parameters and produce predicted distribution maps. One such study was completed by Jaberg and Guisan (2001) to model the distribution of bats in relation to landscape structure. Again, the authors in this study used GLMs to predict the occurrence of bats in areas where no presence/absence data exists. They used presence/absence points from bat surveys and related these observations to a digital elevation model (DEM) of the area and its derived products. In addition, they created land-cover classes by using the percent coverage of certain landscape elements. The authors found that correlations between the

observed species regularity and predicted values were fair to good and the overall prediction success of the models was 71% to 87%. In addition they found that, although their models were fairly successful for these highly mobile species, model stability and agreement tended to decrease with increases in the mobility of the species (Jaberg and Guisan, 2001). However, their models were limited to large scale variations in landscape features because their DEMs were gridded at 25m and 100m resolution. Although this technique of using coarse resolution grids may work for species such as bats, many species, such as rockfish, are associated with features that can only be delineated on a finer scale.

Compared to the amount of modeling studies used to predict the occurrence of terrestrial species, very few studies have been completed on marine organisms. One study combined multibeam and backscatter data with video data to classify habitats and associate fish with those habitats on Heceta Bank, Oregon. The main finding from this study was that sonar can be used along with submersible data to allow for habitat-based stock assessments of groundfish species (Nasby-Lucas et al., 2002). The authors were able to predict the abundances of several species of groundfish by breaking the area into defined habitat patches and using the density of fish from the video data to determine the density of fish throughout that habitat. However, this study did have some shortcomings. First, the substrate types were visually interpreted. Although this is often a useful technique for distinguishing between different habitats, it is time consuming and very subjective. Second, they did not have extensive video data that covered all the representative habitats on Heceta Bank. Therefore, it was difficult to extrapolate their findings to the entire bank. Finally, they did not have enough video data to test their density predictions. They could not take a subset of their video to determine the accuracy of their results because all the video footage was used in making their predictions (Nasby-Lucas et al., 2002).

A similar study that was able to successfully predict habitat suitability was recently completed by Iampietro et al. (2005). Here the authors were able to create predictive models of habitat suitability and fish distribution using analyses on a high resolution multibeam bathymetry DEM from the Del Monte Shale Beds in Monterey Bay, California. These data were then used to make stock estimates for the study area based on topographic position index (TPI). Although successful in the ability of their model to accurately predict the habitat suitability for approximately 80% of the rockfish observed in ROV video transects, this study was conducted across a relatively shallow depth range (15-75m) and over relatively low habitat relief (<2m).

One of the shortcomings of this study was that the resolution of the DEM (1-2m) was very close to the magnitude of the relief and could not support slope or rugosity analyses which could have further improved the accuracy of the models beyond the use of TPI alone. Also, the models created by Iampietro et al. lacked statistical rigor; parameters and weighting factors used in their simple additive models were chosen by the researchers based on their observations and trial-and-error experimentation, rather than being fitted by a predictive statistical modeling technique such as GLM (Iampietro et al., 2005).

Therefore, the purpose of this capstone is to determine how well the approach taken by Iampietro et al. (2005) in shallow, low relief rockfish habitats can be applied to deeper, high relief rockfish habitats where the use of slope and rugosity models can be supported in addition to TPI. The study site for this project, Cordell Bank, is a granitic bank located 22 miles offshore from Pt. Reyes, California. Previously part of the Sierra Nevada Mountains about 93 million years ago, Cordell Bank was sheered off the North American plate by the Pacific Plate and carried north until it reached its current location. Between 20,000 and 15,000 years ago, when sea level was 120 meters below its current level, the bank was a true island (Cordell Bank National Marine Sanctuary, 2007). Now submerged, Cordell Bank offers an ideal location to conduct this study because it has a large depth range (40-200m), highly varied geomorphic relief (1-15m), and a substantial amount of biological video data from submersible transects (Figure 1). The combination of these characteristics allowed me to build upon the methods utilized by Iampietro et al. (2005).
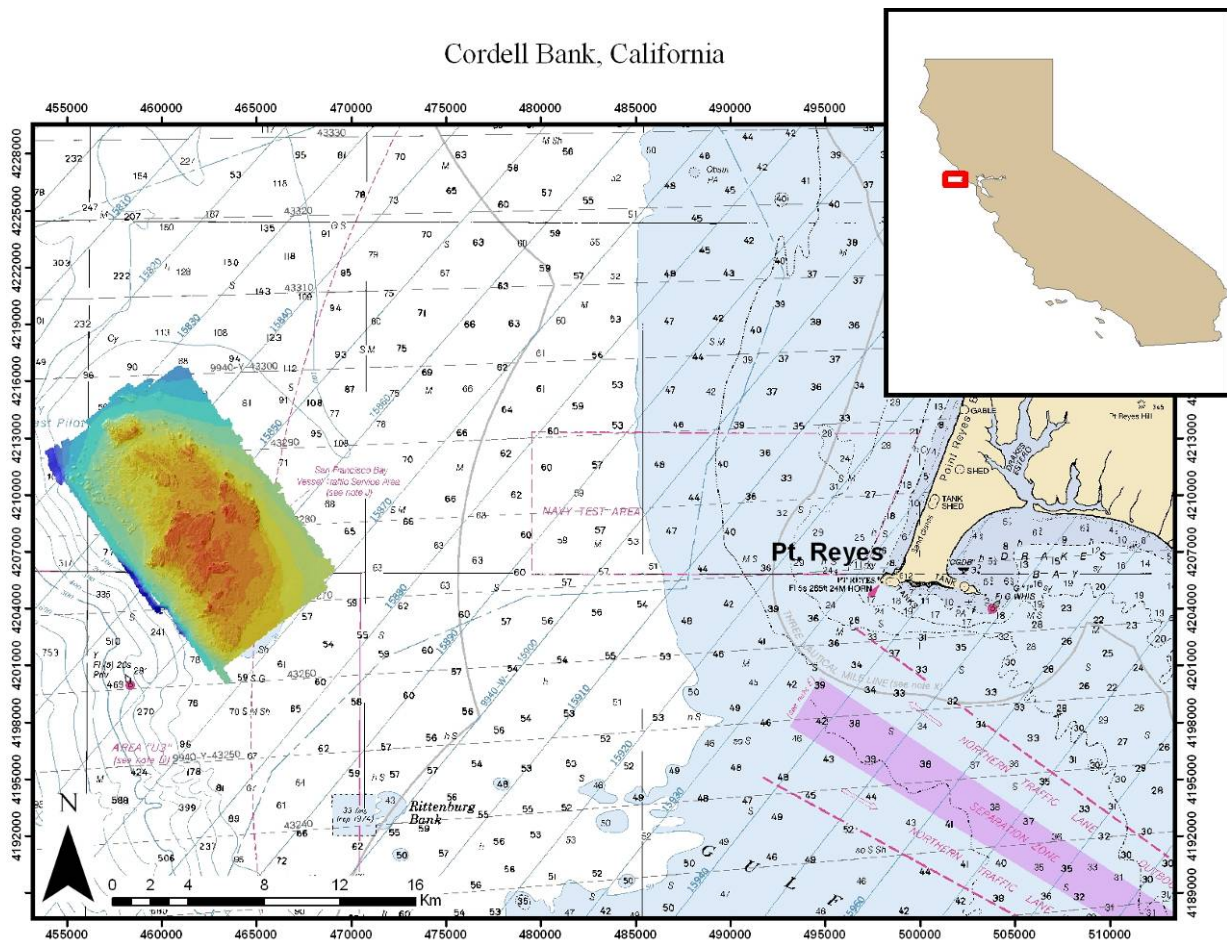
Figure 1: Location of Cordell Bank in relation to Pt. Reyes, California. Multibeam bathymetry digital elevation model of Cordell Bank shown in shaded relief and colored by depth, with red being the shoalest soundings. (Image resolution: 3m; Coordinate System: UTM 10N WGS 84; NOAA Chart 18640, soundings in fathoms; Multibeam survey completed fall 2005.)

This study addressed the following questions:

1) What seafloor habitat characteristics of Cordell Bank can be derived from high-resolution multibeam sonar data?

2) Can GIS landscape ecology modeling techniques make use of these sonar-derived habitat characteristics to accurately predict the distribution and stock sizes of three different species of rockfish: yellowtail (*Sebastes flavidus*), rosy (*S. rosaceus*), and greenstriped (*S. elongatus*)?

   H_0: There is no relationship between the sonar-derived habitat characteristics of Cordell Bank and the distribution of rockfish.

H$_1$: GIS Landscape ecology modeling techniques and multibeam bathymetry data can be used to more accurately predict the distribution of rockfish than submersible video transect survey techniques alone.

If the alternative hypothesis proves to be correct and this is an effective technique, it could be expanded and tested over broader regions of the California coast, particularly in the design and placement of management areas for groundfish species. The general approach was to use the CSUMB Seafloor Mapping Lab multibeam and sidescan data collected in fall of 2005 combined with submersible video data collected in 2002 by Cordell Bank National Marine Sanctuary to create habitat maps within GIS and attempt to associate the different species with their habitats. Habitat parameters including depth, slope, rugosity, topographic position index (TPI), and seafloor substrate type were then used to delineate the different habitats.

2 METHODS

Site Description

In order to complete this study, both multibeam and backscatter data were collected by the Seafloor Mapping Lab at California State University, Monterey Bay aboard the R/V *VenTresca* at Cordell Bank National Marine Sanctuary, California. Cordell Bank is a granitic bank located 40 km west of Point Reyes, California and is home to an array of marine organisms. Often considered a submerged island, Cordell Bank is 4.5 miles wide by 9.5 miles long with pinnacles coming to within 40m of the water's surface. The upwelling provided by the California Current makes this a prime location for a number of resident and migratory species. From the resident species, rockfishes (*Sebastes spp*.) are by far the most abundant fishes, making up 90% of the fish observations on the bank. Although the sanctuary status of Cordell Bank does not protect them from fishing pressure, it does appear to provide a natural refuge for overfished species such as boccacio (*S. paucispinis*), yelloweye (*S. rubberimus*), vermilion (*S. miniatas*), and canary rockfish (*S. pinniger*) within its diverse habitats (Pirtle, 2005).

Multibeam Bathymetric Sonar Data Collection and Processing

Multibeam bathymetric sonar is collected by using sound beams to ensonify a large swath of the seafloor. One transmit beam, which is wide across-track (>150˚) and narrow along-track

(1.5˚), is released from the sonar head at equally timed intervals as the vessel travels through the water. Once the beam bounces off of the seafloor and travels back up to the sonar head, it is intercepted by the receive array and broken up into 101 1.5˚ x 1.5˚ individual "beams." Each beam is associated with a two-way travel time and a velocity. This gives depth data across the entire 150º swath to create topographic maps of the seafloor (Figure 2). These maps can be used to derive characteristics that allow for the delineation of habitats such as slope, rugosity, and topographic position index.
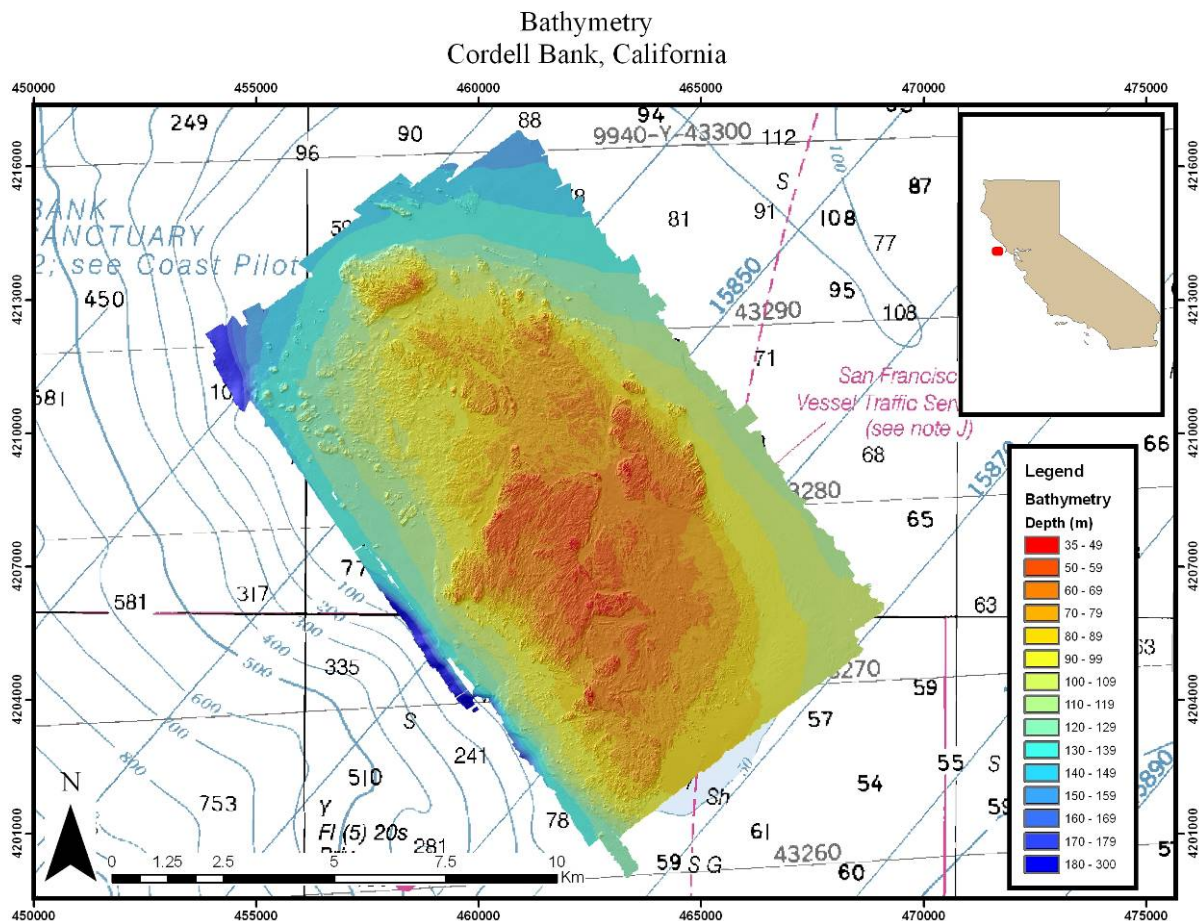
Figure 2: Three-meter resolution digital elevation model (DEM) of Cordell Bank shown in shaded relief and colored by depth produced from the multibeam data collected by the Seafloor Mapping Lab at California State University, Monterey Bay. (Image resolution: 3m; Coordinate System: UTM 10N WGS 84; NOAA Chart 18640, soundings in fathoms; Multibeam survey completed fall 2005.)

The multibeam and backscatter data for this project were collected over 10 days in September and October of 2005 at Cordell Bank National Marine Sanctuary located 40km west of Pt. Reyes. To collect the data, a pole-mounted Reson SeaBat 8101 sonar head with 101 1.5°

beams was used. During collection, the data were collected along with corrections for pitch, roll, heave, and heading using the Applanix POS/MV heading and motion sensor with ±0.02º accuracy. In addition, the position was provided by a C-Nav GPS receiver mounted on the boat. All these data were captured using a Triton Imaging, Inc. Isis Sonar data acquisition system, which also displayed the data in real-time through a program called DelphMap.

After the collection was complete, the data were brought back into the lab for post-processing. The multibeam data were imported into Caris HIPS software where they were corrected for errors in attitude, tide, and sound velocity corrected with data collected by a sound velocity profiler. Then they were cleaned using standard multibeam processing procedures in Caris HIPS. Following cleaning, the data were exported as XYZ (Easting, Northing, depth) data from the bathymetry associated with statistical error (BASE) surface created within Caris. This process, exports XYZ point data at regularly-spaced intervals equal to the resolution of the BASE surface. Once exported, they were brought into a program called Fledermaus for quality control (QC) to verify that there were no spikes or erroneous data points remaining. Subsequent to QC, the data were exported as an ArcView Grid for GIS analysis.

After processing the data, the cleaned datasets were brought into ArcGIS for analysis. First, the ArcView grid exported from Fledermaus was brought into ArcGIS where it was projected as a digital elevation model (DEM). A DEM is a raster dataset that consists of elevation values at regularly spaced intervals. Since many species of rockfish prefer rocky areas with high relief or areas of large boulders and stones, analyses were performed on the DEM to delineate these different habitats.

2.1 *Slope Analysis*

The first type of analysis used was a slope analysis. Slope is calculated by determining the max slope value between an individual DEM cell and its eight neighbors. Slope analysis was completed in ArcGIS using the Spatial Analyst tool to create a new raster dataset, which includes the slope value for each cell in the DEM. These slope values were then broken up into the slope categories from the deep water marine benthic habitat classification scheme (Greene et al., 1999) (Table 1). Slope could possibly be a good predictor of rockfish abundance because areas with high slope values are often associated with high relief.

Table1: Seafloor slope categories based on the deep water marine benthic habitat classification scheme (Greene et al., 1999).

Slope	Category
0-1°	Flat
1-30°	Sloping
30-60°	Steeply Sloping
60-90°	Vertical
>90°	Overhang

2.2 *Multibeam Substrate Analysis*

The second type of analysis used was a rugosity analysis. Rugosity is the ratio of surface area to planar area of the terrain, which can be used to measure the roughness of the seafloor. Flat, smooth areas have rugosity values near 1 while higher values are associated with areas of higher relief or bumpiness. This could be a good predictor of rockfish abundance because they tend to situate themselves in crevices or between rocks. Where rugosity is high, there is a greater amount of potential habitat space for the rockfish. Rugosity analysis calculates the surface area ratio for each cell in a DEM grid using the elevation of the cell and its eight neighbors. Rugosity grids were created using the Benthic Terrain Modeler (BTM) extension for ArcGIS. The BTM was created by Davey Jones' Locker Seafloor Mapping and Marine GIS Laboratory, Department of Geosciences at Oregon State University, and the National Oceanic and Atmospheric Administration (NOAA) Coastal Services Center (Wright et al., 2005). The rugosity tool in the BTM is an adaptation of the Surface Areas and Ratios extension for ArcView 3.x (Jenness Enterprises, 2005).

From the rugosity analysis, a binary raster dataset was created that only includes two values: 0 for soft sediment and 1 for rocky substrate. The breakpoints for these two classes were rugosity values of 1-1.05 for soft sediment and >1.05 for hard substrate (rock). To determine how well the class breaks matched up with a visual interpretation of the substrate type, 100 random points were placed throughout the grid and were visually interpreted and attributed with values of 0 or 1 for soft sediment or hard substrate, respectively. The visual interpretation was then compared with the supervised classification from the rugosity grid and an accuracy assessment was done by determining the percentage of points that were classified accurately by

9

dividing the number of points where the results were the same for both the automated and visual classifications by the total number of accuracy assessment points.

2.3 *Topographic Position Index (TPI) Analysis*

Topographic position index (TPI) is a measure of relative elevation, which indicates the position of a given point in the overall surrounding landscape. TPI can be used to identify and delineate landforms such as peaks, ridges, cliffs, slopes, flat plains, and valleys, and is calculated by comparing the elevation of each cell in a DEM to that of its surroundings. Because the neighborhood size of the surroundings used for the elevation comparison can be adjusted, TPI can be calculated at various scales. Thus, an analysis size can be chosen that will identify features of any desired size, ranging from small-scale features such as the tops of boulders and pinnacles, to entire reefs, to regional-scale features such as seamounts (all classified as "peaks" at different scales). Likewise, TPI can be used to locate fissures and cracks in rock, sand channels, and submarine canyons (TPI "valleys" of increasing scale). In fact, the only limiting factor is the resolution (cell-size) of the DEM, which determines the minimum scale of features that can be delineated.

The TPI analysis employed in this study was done using the algorithm of Weiss (2001), which uses an annulus- ("donut") shaped neighborhood. TPI is calculated using the formula:

tpi<scalefactor> = int((dem - focalmean(dem, annulus, irad, orad)) + .5)
where:
scalefactor = outer radius in map units
irad = inner radius of annulus in cells
orad = outer radius of annulus in cells

The scale of the analysis is defined by the inner (irad) and outer (orad) radii of the annulus neighborhood. The results of this calculation can range from strongly positive (areas that are higher than their surroundings at the specified scale), to strongly negative (areas that are much lower than their surroundings). Intermediate values define irregularly sloping areas, while flat areas and areas of constant slope result in values near zero. Because the magnitude and range of the results are DEM-specific, the initial TPI cell values are classified into standard deviation

classes, which are categorized into classes such as "peak", "slope", and "valley". In order to differentiate the ambiguous near-zero TPI values into "slope" and "flat" classes, the previously mentioned slope grid derived from the DEM is used, with a slope value of 5° serving as the break point between the two classes.

TPI products were created using the bathymetric position index (BPI) grid creation and classification tools included in the BTM to create grids at a variety of scales ranging from 30m to 240m (orad), with a 5-cell (15m) annulus thickness (orad - irad). The Standardize BPI Grids tool was then used to standardize the outputs from the raw BPI analysis. Since the BPI values for the broad scale and fine scale grids differ due to the spatial autocorrelation of the data (i.e. locations that are closer together are more related than locations that are farther apart), the range of BPI values increases with scale. Therefore, the standardization of the BPI grid allows for the classification of BPI datasets at differing scales (Weiss, 2001). The Standardize BPI tool utilizes the following algorithm within the Raster Map Algebra Operation object available through ArcGIS Spatial Analyst to standardize TPI grids by converting them to z-scores:

TPI<scalefactor>_s = int((((TPI<scalefactor> - mean) / std dev) * 100) + 0.05)

scalefactor = outer radius in map units * input bathymetric data set resolution(cell size)
mean = mean cell value across TPI data set
std dev= standard deviation of cell values across TPI data set

This process was used to create standardized grids for each scale of raw TPI grid. The resulting standardized grids have x = 0 and SD = 100.

Once the TPI grids were standardized, slope position grids were created by classifying the standardized grid cell values into standard deviation classes with 0.5 standard deviation increments (<-1.00, -1.00 to -0.50, -0.50 to 0.50, 0.50 to 1.00, and >1.00) using the reclassification tool within Spatial Analyst to assign each class a value from 1 to 6, respectively, excluding 4. Next, the Map Algebra tool was utilized along with the slope grid to split the ±0.5 standard deviation class into flat (slope < 5°, class 3) and slope (slope >= 5°, class 4) classes. This process produced a slope position grid with six classes that were assigned the following

landscape feature values: 1 = Valley/Crevice, 2 = Lower Slope, 3 = Flat/Plain, 4 = Middle Slope, 5 = Upper Slope, and 6 = Peak/Ridge.

Acoustic Backscatter Data Collection and Processing

Acoustic backscatter data (also known as sidescan data) are collected using sonar, but rather than measuring depth, the intensity of the return echo is recorded. This allows for the classification of different substrate types due to the differences in amplitude of the returning sound. For example, the sound returning from a rock would have greater amplitude than the sound returning from soft sediments which tend to absorb more and reflect less of the sonar's transmitted sound. In addition, the orientation and texture of the substrate also affect the amplitude of the returning sound. The backscatter data can be converted into a georeferenced sonograph mosaic for use in the visual interpretation or supervised classification of the surface texture of the seafloor.

Both the multibeam data and the acoustic backscatter data were processed and used for habitat analysis. The same raw XTF files used for the multibeam were brought into Isis Sonar for sidescan processing. Within Isis, each track line was individually mosaicked and exported and brought into DelphMap where they were exported as geotiffs. Next, the data from each individual survey track line were brought into a GIS program called TNTmips where the usable data from each line was extracted and put into a chosen order to create an acoustic backscatter mosaic of the entire site. The acoustic backscatter mosaic was exported as a geotiff image for analysis (Figure 3).
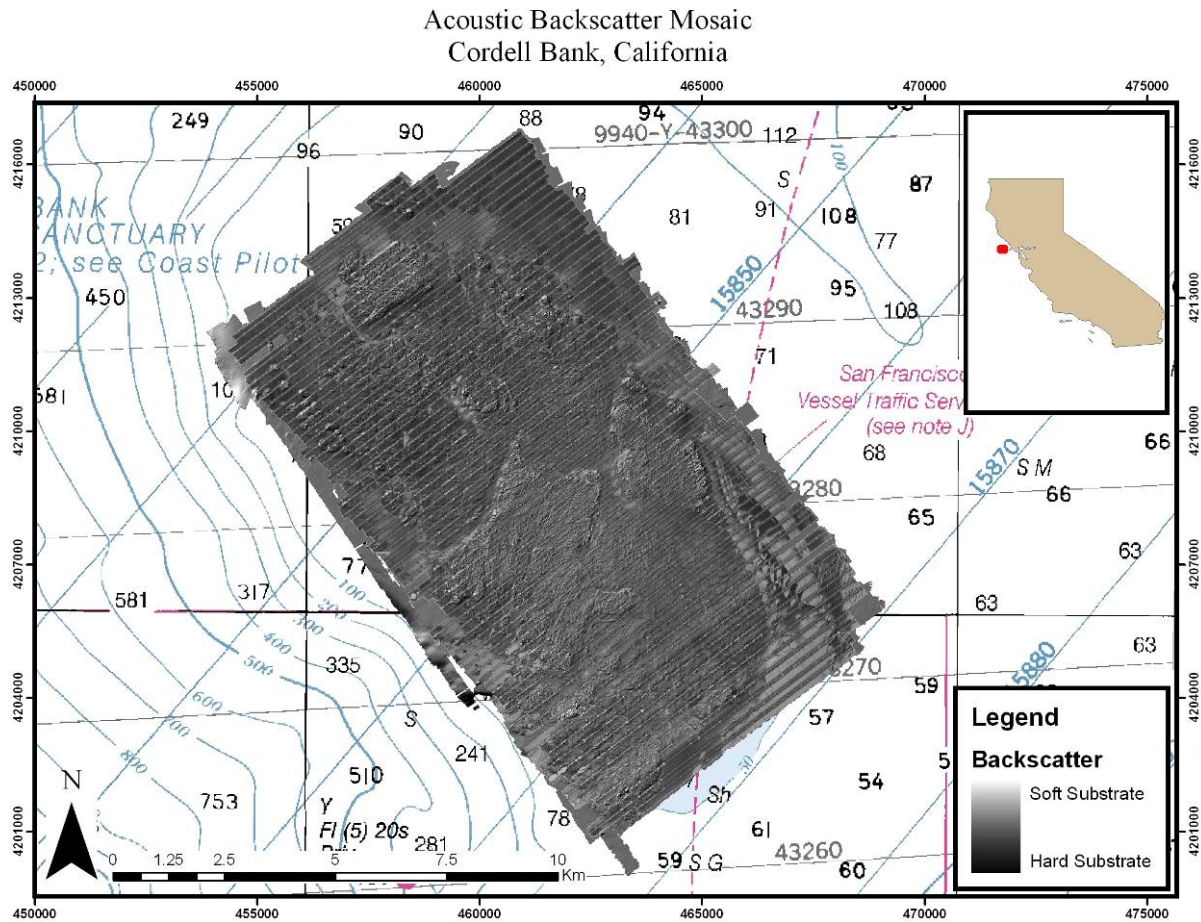
Figure 3: Mosaic created from the acoustic backscatter data collected at Cordell Bank. The darker areas are rocks or coarser sediment while the lighter areas are mud or finer sand. (Image resolution: 0.5m; Coordinate System: UTM 10N WGS 84; NOAA Chart 18640, soundings in fathoms; Survey completed fall 2005.)

2.4 *Acoustic Backscatter Analysis and Habitat Classification*

After processing the backscatter data, the geotiff image created in TNTmips was brought into custom designed software to calculate homogeneity and entropy and output geotiffs with the calculated values (Cochrane and Lafferty, 2002). These values are calculated by looking at all the pixels in an 11 x 11 co-occurrence matrix and assessing the spatial relationship of pixel intensities. Entropy grids were created by determining the degree of difference between the pixel intensities while homogeneity grids were created by determining the degree of similarity between the pixel intensities (Intelmann and Cochrane, 2006).

Once created, the entropy and homogeneity grids were layered with the backscatter mosaic inside ArcGIS for the completion of a supervised classification of substrate type. The

13

images were classified into three categories of bottom type using the maximum likelihood classification (MLC) tool in Spatial Analyst: soft, mixed, and hard bottom. The MLC works by determining the range of pixel values from the different signatures and using those pixel values to classify different sediment types throughout the image. Once good signatures were chosen for the MLC and a relatively clean image was created, the image was cleaned further using the Raster Painting tool within spatial analyst. Since no data is collected directly beneath the sonar head (at nadir) when acquiring sidescan data, 'stripes' end up in the final image. During the MLC, much of this 'data' was misclassified into the rock and mixed categories. In order to edit the misclassified data at nadir, a Euclidean Distance Grid was created and a new MLC was performed using signatures placed within the misclassified data at nadir. This process created a new grid that was reclassified and utilized to create a polygon shapefile to extract out the misclassified values at nadir. The Block Statistics Majority tool with Spatial Analyst was used to interpolate more accurate values into the areas surrounding nadir and these values were merged with the original grid. Then, six video transects were used to verify the accuracy of this classification by attributing points along each transect with the type of substrate identified in the video and then comparing the visual interpretation to the supervised classification of the acoustic backscatter.

Once a clean image was created, the classification was converted to a polygon shapefile. Attributes were then added to each of the polygons based on the Greene et al. (1999) deep-water marine benthic habitat classification scheme. These attributes included the following: megahabitat, seafloor induration, seafloor slope, seafloor complexity, depth, macro/micro habitat, and geologic units.

2.5 *Rockfish/Habitat Associations Analysis Using GLMs*

For this analysis, three species of rockfish were chosen: *S. flavidus, S. elongatus,* and *S. rosaceus*. Because the original fish observation data from submersible video analysis consisted of point locations with multiple fish counts by species, these data were converted to presence/absence format for use in binomial logistic regression models. To accomplish this, the Random Point Generator extension for ArcView 3.x (v1.3, Jenness Enterprises, 2005) was used to "explode" the single location for each fish observation into *n* randomly-located points within a 10m radius of the original location, where *n* = the number of fish of the species of interest

observed at that location. A 10m radius was chosen based on the effective visibility during the submersible dives. Thus, each multiple-fish observation point was converted into many single-fish "presence" points. To create point locations for fish "absence", the submersible transect navigation data were randomly subsampled after excluding all locations where fish were observed. In this manner, a series of point locations were created where the absence of fish was confirmed by direct observation. All fish presence and absence point locations were created in the ESRI shapefile format.

The presence and absence point location shapefiles were used together with the DEM and derived habitat parameter rasters in ArcGIS to create predictive models using the ArcRstats toolbox (Best et al., 2005). ArcRStats integrates ArcGIS with the R statistical package to produce multivariate habitat prediction rasters. Within this toolbox, the generalized linear model (GLM) tool was utilized to produce predictive grids for each species of rockfish based on their associations with the habitat on which they were found. GLMs are similar to linear models but they are believed to be better for analyzing ecological relationships because they do not force the data into unnatural scales, which would be required for a linear model. GLMs require the data to be neither linear nor have constant variance and they are capable of using data from a number of different probability distributions (Guisan et al., 2002). The generalized linear model within ArcRStats works by sampling the values of each of the predictor rasters (i.e. slope, rugosity, depth, etc.) underlying the presence and absence points locations and then using those data to create prediction rasters that display the probability of species occurrence for unsampled locations (Guisan et al., 2002). A binomial logistic regression model is used.

Due to the complexity of the GLM analysis and the extremely large data arrays required, computer memory addressing limitations (4 GB in a 32-bit OS) required subdivision of the Cordell Bank datasets into six smaller areas. The bank was broken up systematically by dividing the bank into six relatively equal parts (Figure 4).
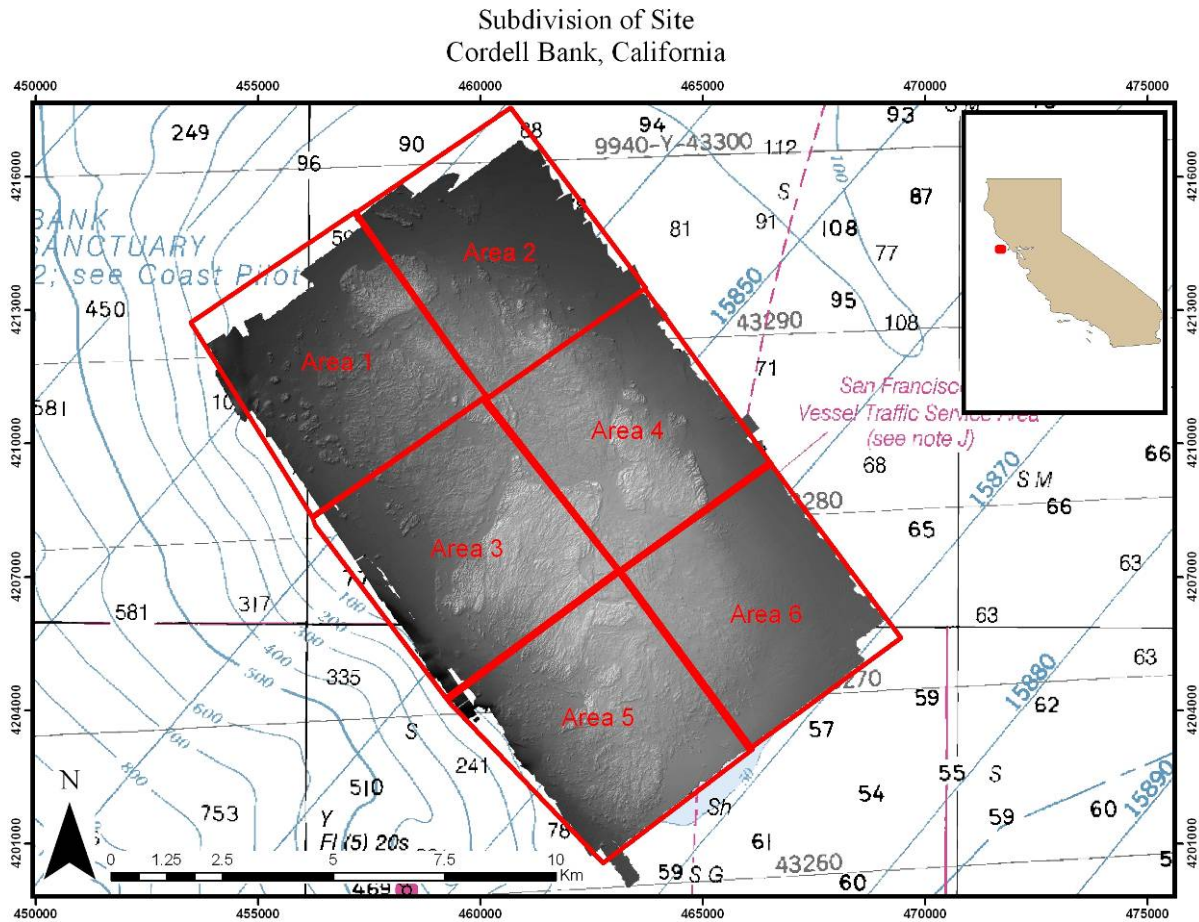
Figure 4: Map displaying Cordell Bank subdivided into six blocks. Each separate block is outlined and labeled in red. (Image resolution: 3m; Coordinate System: UTM 10N WGS 84; NOAA Chart 18640, soundings in fathoms; Multibeam survey completed fall 2005.)

The GLM analyses were performed using the fish presence/absence and habitat raster data for each block separately. However, the same general methods were applied to each block. Each time the ArcRstats GLM tool was used, the appropriate fish presence/absence point locations were specified along with the habitat parameter rasters of interest. These rasters included the bathymetric DEM and derived parameters such as aspect, slope, rugosity, the substrate grid from the rugosity analysis, TPI at the broad and fine scale levels, as well as substrate classification from the acoustic backscatter analysis. Because results from preliminary models suggested that rugosity and slope were correlated, these variables were not used together in any of the models. Separate models containing either rugosity or slope were tested to determine which model was a more effective predictor of fish distribution in each case was and the best model was chosen for analysis. If any of the predictor variables included in the model

were not significant in the model, those variables were removed and the model was re-run without them. The GLM tool model outputs include a prediction grid, a summary of the analysis, plots showing the correlations being paired predictor variables used in the model, an R workspace containing all of the data used in the model, and histograms for each of the input rasters showing the distribution of the presence and absence points vs. the values of each raster.

To test model significance, a D^2 value, which is equivalent to the R^2 in linear models, was calculated by using the following equation:

D^2 = (Null Deviance – Residual Deviance)/ Null Deviance

Where:

Null deviance = the deviance of the model with the intercept only and

Residual deviance = the deviance that remains unexplained by the model after the variables have been included.

Since a perfect model would have no residual deviance (residual deviance = 0), a D^2 value close to one is an indication of good model fit. Since each of the models includes a different number of observations and a different number of predictors, those values have to be taken into account for the D^2 value to be representative of the real fit. Therefore, the equation for the adjusted D^2 value for GLMs is:

$D^2 = 1 – [(n - 1)/(n - p)] \times [1 - D^2]$

Where:

n = the number of observations in the model and

p = the number of parameters in the model

Therefore, the D^2 value increases with an increasing number of observations or a decreasing number of parameters and provides an indication about the performance of the model.

Finally, in order to test how well the model predicts the location of each species of rockfish, the reserved fish presence/absence point locations were plotted on top of the model results raster. The probability values from the predictive model raster were sampled at each reserved fish presence/absence point location. Then, a 2x2 contingency table was created that displayed the proportion of the evaluation points that were correctly classified (i.e. presence points that fell into areas of high probabilities of occurrence and absence points that fell into areas of low probabilities of occurrence). For the contingency table analysis, a threshold value of 0.4 was used as an "expected" value for low probability of occurrence fish (absence), while

0.6 was considered high probability of occurrence (presence). These values were chosen because many modeling studies use a threshold of 0.5 and, by increasing the threshold to 0.6, the chance of predicting rockfish where there were none observed is decreased.

2.6 *Submersible Video Data Analysis*

In order to determine the locations and counts of the different species of fish on Cordell Bank, submersible video data was used. This video data was collected in the fall of 2002 by researchers aboard the *Delta* submersible supported by the *R/V Velero IV*. The personnel for this project included collaboration between the *Delta* crew and scientists from Cordell Bank National Marine Sanctuary (CBNMS), the National Marine Fisheries Service Santa Cruz laboratory, the California Department of Fish and Game, and Washington State University. Between September 22nd and October 3rd, a total of 62, 15 minute transects were completed in 28 dives. Most of the transects were concentrated on the bank in depths of 50 to 100m with some on the continental shelf (80-150m) and continental slope (180-350m) (Cordell Bank National Marine Sanctuary Cruise Report, 2002).

The dive videos were then used to quantify fish. J. Baltan and L. Snook assigned fish observations, including the identification, abundance, and maximize size of the fish, to individual times along the transect. These times were then associated with the coordinates of the *Delta* submersible to give location to the fish observations. The rockfish species (*Sebastes* spp.) most commonly identified on the bank included yellowtail (*S. flavidus*), widow (*S. entomelas*), rosy (*S. rosaceus*), rosethorn (*S. helvomaculatus*), swordspine (*S. ensifer*), starry (*S. constellatus*) boccacio (*S. paucispinis*), canary (*S. pinniger*), yelloweye (*S. ruberrimus*), vermillion (*S. miniatus*), and greenstriped rockfish (*S. elongatus*).

The following three species of rockfish observed within the video data were chosen for the habitat association analysis: *S. flavidus, S. rosaceus,* and S. *elongatus*. These three species were chosen due the number of observations on the bank and the different types of habitats that they are associated with. Both *S. flavidus* and *S. rosaceus* are often associated with rocky habitat. *S. flavidus*, however, tend to be found on high relief habitat such as boulders and sheer rock walls while *S. rosaceus* are found over a broader range of habitat types from areas with low-lying rocks and sand to high-relief rock. On the other hand, *S. elongatus* are mainly associated with low-relief habitats including cobble, rock rubble, and mud (Love et al., 2002).

From this processed data, figures were created that show the distribution of the three species of interest along depth and habitat gradients. First, each fish observation was assigned to a depth class based off of the 3m resolution bathymetry DEM of Cordell Bank. Then the total area observed in each of those depth classes was calculated by creating a 10m buffer around the points in each of the depth classes and calculating the area of the buffer. Ten meters was chosen for the buffer because the average visibility during the dives was 10m. It was assumed that the fish counts included all the fish within the visibility range of the submersible. Once the fish data were attributed with a depth class, the attribute table was exported from ArcGIS and brought into Excel for further analysis. The data was broken up by species and then further into the different depth classes. A density value was then calculated for each species in each depth class by dividing the total number of fish counted in that depth class by the total area that video data was collected from within that depth class. The density of fish was then graphed by depth for each species.

Jodi Pirtle, a Masters student from Washington State University, previously characterized the habitat data from the video transects. Therefore, her habitat classification was utilized for the video analysis. She assigned habitat type based on a two-letter code using eight different categories of geological substrate (Table 2). The primary code represents the substrate type that accounts for between 50% and 80% of the habitat patch. The secondary code accounts for the substrate type that makes up 20 - 50% of the patch (Pirtle, 2005).

Table 2: One-letter code and its associated habitat type from the habitat analysis done on Cordell Bank based on the video data (Pirtle, 2005).

One-Letter Code	Substrate Type
R	Rock Ridge (high to low relief)
B	Boulder (high to low relief)
C	Cobble (low relief)
P	Pebble (low relief)
G	Gravel (low relief)
F	Flat Rock (continuous, low relief)
S	Sand Grains (grains distinguishable)
M	Mud (noticeable organic particles)

Prior to analysis, the video transects were separated into two subsets: one subset that would be used to make predictions for the distribution and abundance of the rockfish (training

transects) and one subset that would be used to evaluate the accuracy of the predictions (evaluation transects). Once sub-divided, the training transects were converted from a point shapefile to a polyline shapefile with a separate line for each transect using Hawth's Analysis Tools within ArcGIS (Beyer, 2004). A 10m buffer was created around each of the transect lines to acquire a total area covered by each transect. Then, the percentage of the habitat types within each transect was calculated. This percentage was used to determine the total area of each habitat type within all the training transects. After that, the densities of *S. flavidus, S. rosaceus*, and *S. elongatus* were calculated for every habitat type by dividing the total number of fish in a habitat type by the total area of that habitat type found in the training transects. The densities of the fish within the evaluation transects were also calculated using the same procedures.

2.7 *Comparison of Video Analysis versus GLM Predictions*

To determine the accuracy of the different techniques used to predict the abundance and distribution of fish, a comparison was made between the predictions made from the video analysis and those made from the GLMs. First, the densities of the three species of fish found in each habitat type from the video analysis were used to predict the expected number of fish in the evaluation transects. This was done by multiplying the density of fish per substrate type by the total area of each substrate type in the evaluation transects. This calculation provided a prediction of the number of fish expected to be found in the evaluation transects based on the areas of the different substrate classes. This process was completed on eight representative evaluation transects.

Within the same eight evaluation transects, the total areas of each probability class from the GLM predictions were calculated and multiplied by the density of fish expected to be found in those probability classes. Then, the fish from each probability class were added together to get the total number of fish expected to be found within each evaluation transect. These values were then compared to the results from the video analysis predictions and the actual numbers of fish found in the evaluation transect to calculate a percent error between the predictions and observations for both the results from the video analysis and the GLM predictions. These percent errors were then compared to determine which method was better at predicting rockfish abundance.

2.8 *Stock Assessments*

In order to determine the total stock for *S. flavidus*, *S. rosaceus*, and *S. elongatus*, the results from the GLM analyses were utilized. However, prior to completing the stock assessments, the density of each species within each habitat type had to be calculated. First, the point data from the fish observation shapefile were converted into lines using Hawth's Analysis Tools. A separate line was created for each transect within each dive. Then, the buffer tool within Spatial Analyst was used to produce a 10m buffer around each transect track line. Ten meters was chosen because the average visibility during the submersible dives was an average of 10 meters to either side of the track. The buffer shapefile was used as an extraction mask within each block so that the values from the prediction grids for each species could be extracted out. This produced prediction rasters that only encompassed the area where submersible surveys were conducted. This raster dataset was reclassified into ten classes with increments of 0.1 from 0 – 1. Once reclassified, the raster dataset was converted to a polygon shapefile using the 'Convert Raster to Features' tool within Spatial Analyst with separate polygons for each probability class. The area was then calculated for each polygon and the areas within each class were added up to get a total area of each probability range for all transects in the different areas. After the areas were determined for all the probability ranges, the total numbers of fish within each of those probability classes were counted. Finally, the total number of fish observed from each probability class was divided by the total area of that habitat class within the transects. This result gave an average density of fish that were observed within each probability class.

Once the density of fish in each habitat class was determined, these densities were multiplied by the total area of that habitat class in each area. This process provided a calculated stock size for each species of fish within each area. These values were then combined to determine the stock size of each species of fish over the entire site. The errors associated with these stock counts were also taken into account by multiplying the percentage of error from the original by the stock predictions.

3 RESULTS

3.1 *Multibeam Slope Analysis*

21

The results from the slope analysis shows that the majority of Cordell Bank falls into the "sloping" category (77%) where the degree of slope ranges from 1˚ to 30˚. Areas with a slope of less than 1˚ make up 26% of the bank while the "steeply sloping" and "vertical" classes only make up a very small portion of the bank. There were no slope values that fell within the "overhang" category (Table 3, Figure 5).

Table 3: Summary table of the results from the slope analysis. Each of the categories is listed along with its corresponding area and the percentage of the site it makes up.

Slope Class	Area (m^2)	Area (km^2)	Percentage of Site
1 (Flat 0-1˚)	32515692	32.5	26%
2 (Sloping 1˚-30˚)	95957611	96.0	77%
3 (Steeply Sloping 30˚-60˚)	558225	0.6	0.01%
4 (Vertical 60˚-90˚)	1481	0.0	<0.00%
5 (Overhang >90˚)	0	0	0%
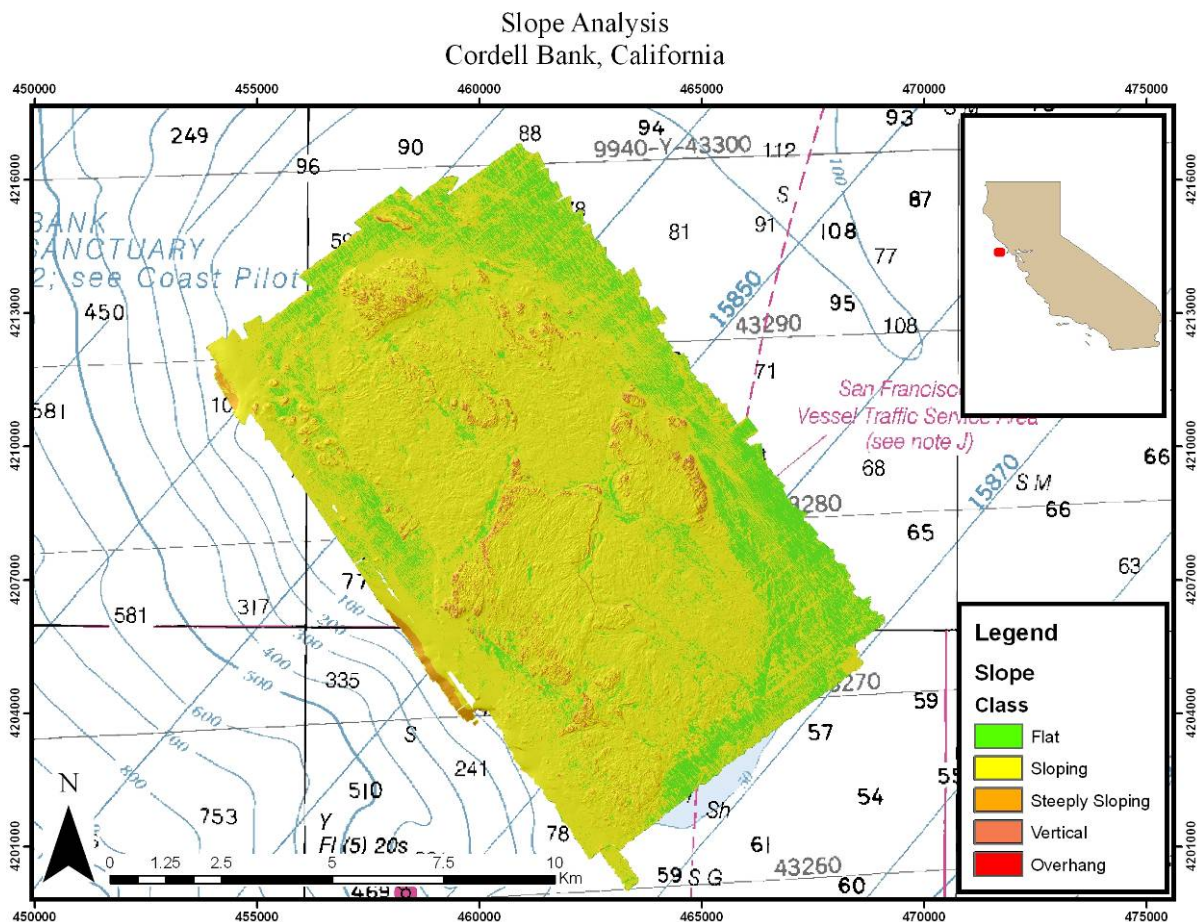
Figure 5: Slope analysis of Cordell Bank, California divided into the seafloor slope categories from the deep water marine benthic habitat classification scheme (Greene et al., 1999)(See Table 3). (Image resolution: 3m; Coordinate System: UTM 10N WGS 84; NOAA Chart 18640, soundings in fathoms; multibeam survey completed fall 2005.)

3.2 *Multibeam Substrate Analysis*

The substrate analysis of the multibeam data from Cordell Bank produced a grid that delineates between hard and soft substrate (Figure 6). In this binary raster, soft substrates were given a value of "0" and hard substrates were given a value of "1." The hard substrate (rock) makes up 36 km^2 (29%) and the soft substrate makes up 88 km^2 (71%) of the survey area. These classified values were compared to the visual interpretation and were found to be 94% accurate.
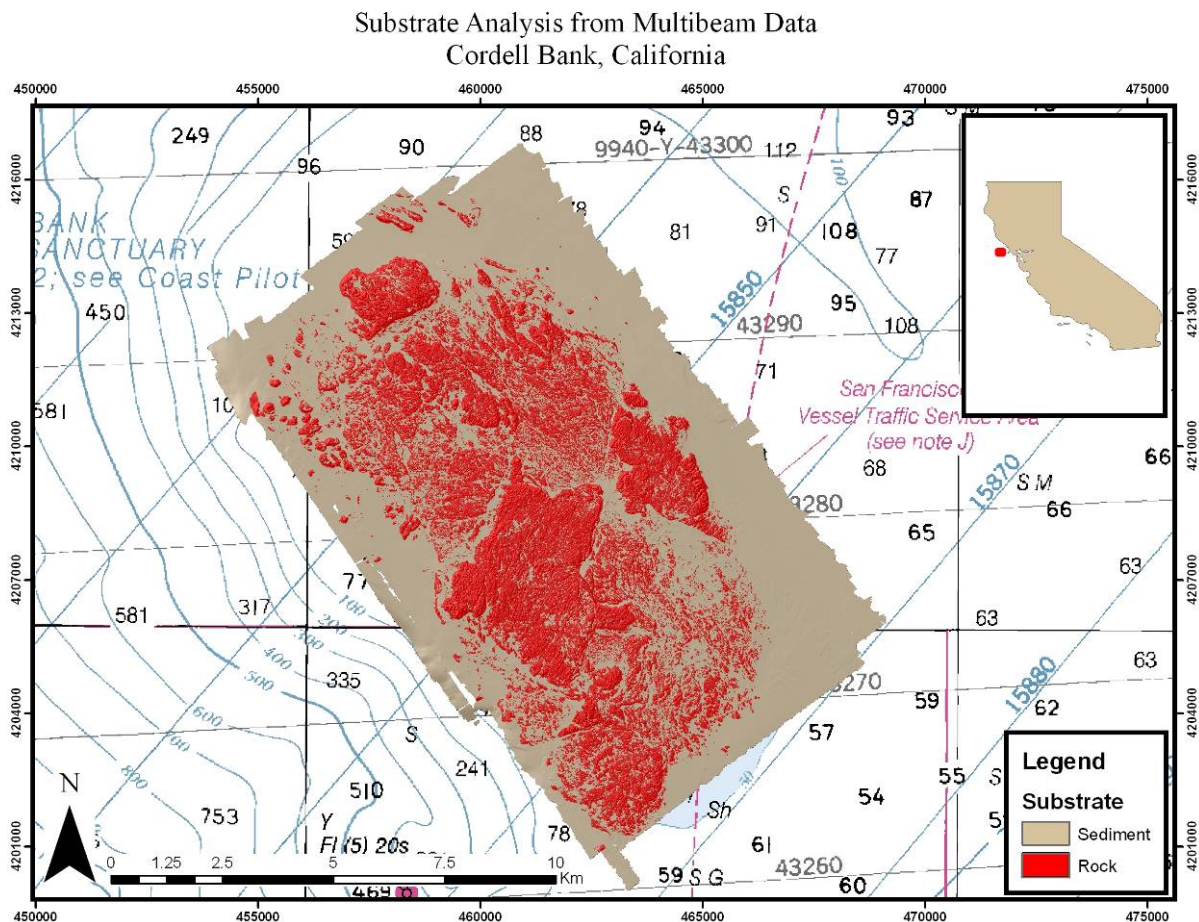
Figure 6: Rugosity (surface area : planar area ratio) values >1.05 were classified as rocky habitat to create a substrate classification grid. (Image resolution: 3m; Coordinate System: UTM 10N WGS 84; NOAA Chart 18640, soundings in fathoms; multibeam survey completed fall 2005.)

3.3 *Topographic Position Analysis (TPI)*

The TPI analysis of Cordell Bank at four different scales (30m, 60m, 120m, and 240m) provided a range of different results (Figure 7). The finer scale analyses provided fewer slope position classes than the broad scale analyses. Using an annulus radius of 30m only provided four separate slope position classes (peak/ridge, middle slope, flat/plain, valley/crevice) (Figure 7), while using an annulus radius of 240m provided all six habitat classes (Figure 8).
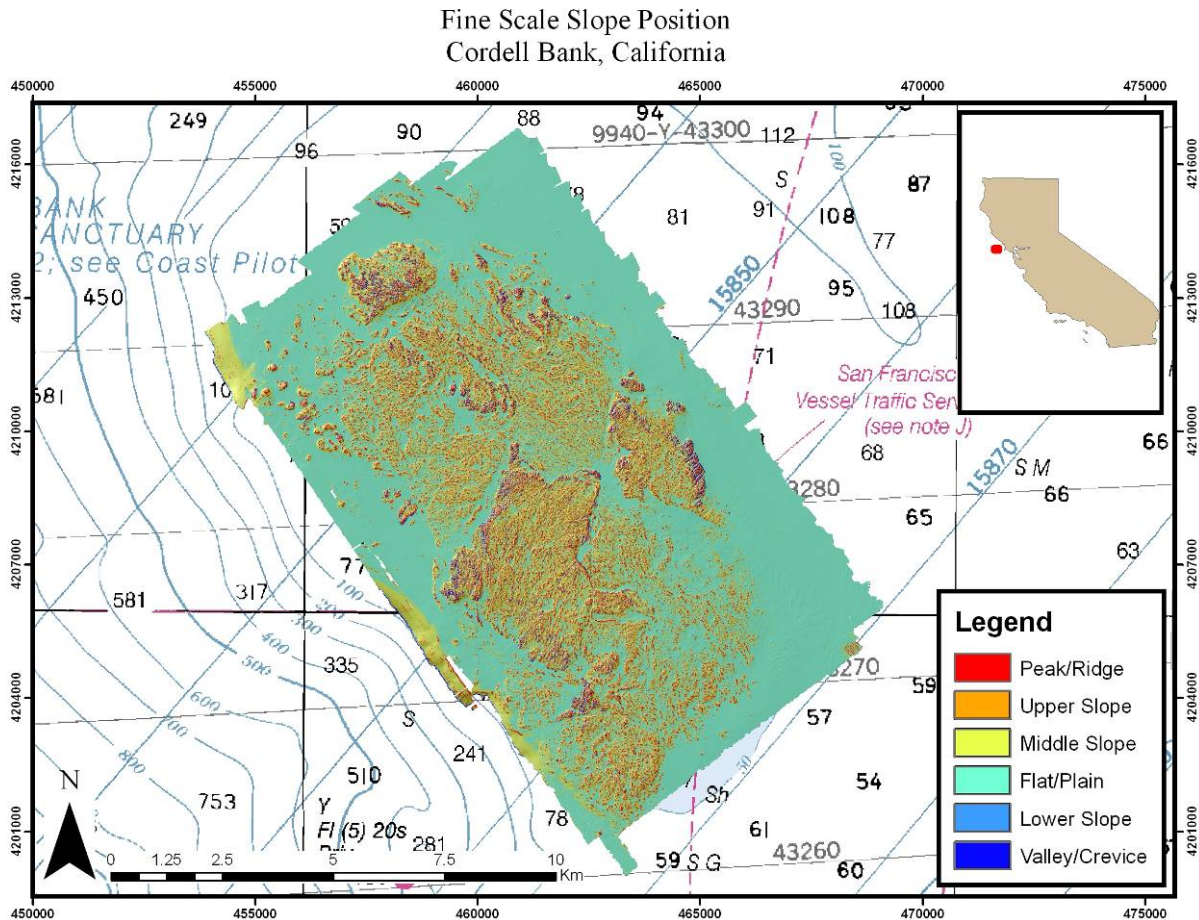
Figure 7: A fine scale (30m annulus) slope position grid derived from a TPI analysis performed on the 3m bathymetry data from Cordell Bank, California. (Image resolution: 3m; Coordinate System: UTM 10N WGS 84; NOAA Chart 18640, soundings in fathoms; multibeam survey completed fall 2005.)

24
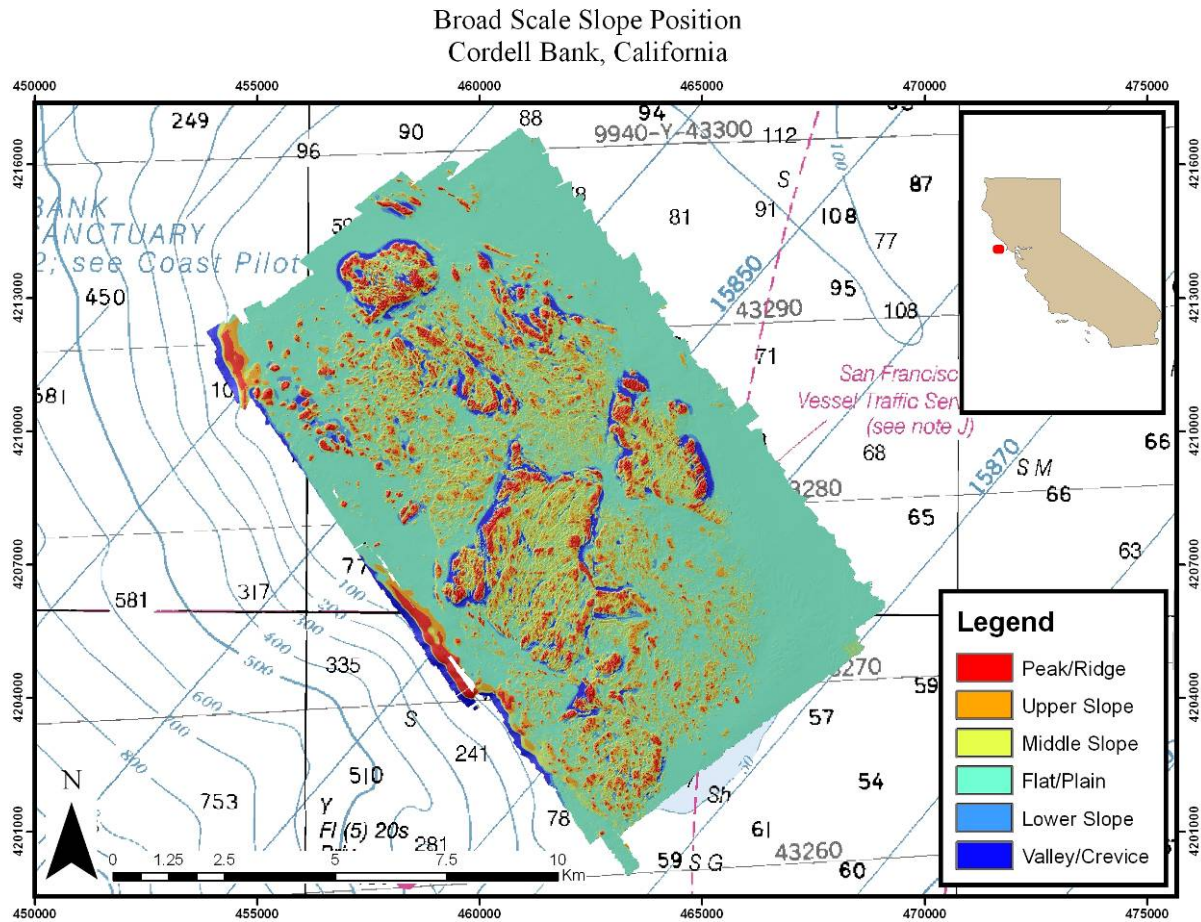
Figure 8: A broad scale (240m annulus) slope position grid derived from a TPI analysis performed on the 3m bathymetry data from Cordell Bank. (Image resolution: 3m; Coordinate System: UTM 10N WGS 84; NOAA Chart 18640, soundings in fathoms; multibeam survey completed fall 2005.)

The most common habitat type across all scales was the "middle slope" class, which took up an area of 60-70% of the site in each analysis. The least common class was the "upper slope" class. This class only made up 4-6% of the bank in the broad scale analyses and did not ever appear in the fine scale analyses (Table 6).

Table 4: Results from the topographic position index (TPI) classification of Cordell Bank. Each habitat class is broken into the four different annulus sizes used to run the classification. The area each habitat class makes up in the different classifications is reported in meters squared, kilometers squared, and the percentage of the site.

Slope Position Class	Annulus Radius(m)	Area (m^2)	Area (km^2)	Percentage of Site
1 Peak/Ridge	30	2447451	2.4	2%
	60	6538113	6.5	5%
	120	4550364	4.6	4%
	240	4437405	4.4	4%
2 Upper Slope	30	0	0.0	0%
	60	0	0.0	0%
	120	7696062	7.7	6%
	240	4810410	4.8	4%
3 Middle Slope	30	84759093	84.8	68%
	60	80098812	80.1	65%
	120	78202386	78.2	63%
	240	77299092	77.3	62%
4 Flat/Plain	30	23022072	23.0	19%
	60	17789562	17.8	14%
	120	21123783	21.1	17%
	240	19181979	19.2	15%
5 Lower Slope	30	0	0.0	0%
	60	11547873	11.5	9%
	120	0	0.0	0%
	240	11129445	11.1	9%
6 Valley/Crevice	30	13916691	13.9	11%
	60	8170947	8.2	7%
	120	12572712	12.6	10%
	240	7286976	7.3	6%

3.4 *Acoustic Backscatter Analysis and Habitat Classification*

The results from the acoustic backscatter analysis produced a polygon shapefile that is attributed with a variety of different habitat characteristics including megahabitat type, substrate type, seafloor slope, seafloor complexity, geologic units, and macro/microhabitats (Figure 9).
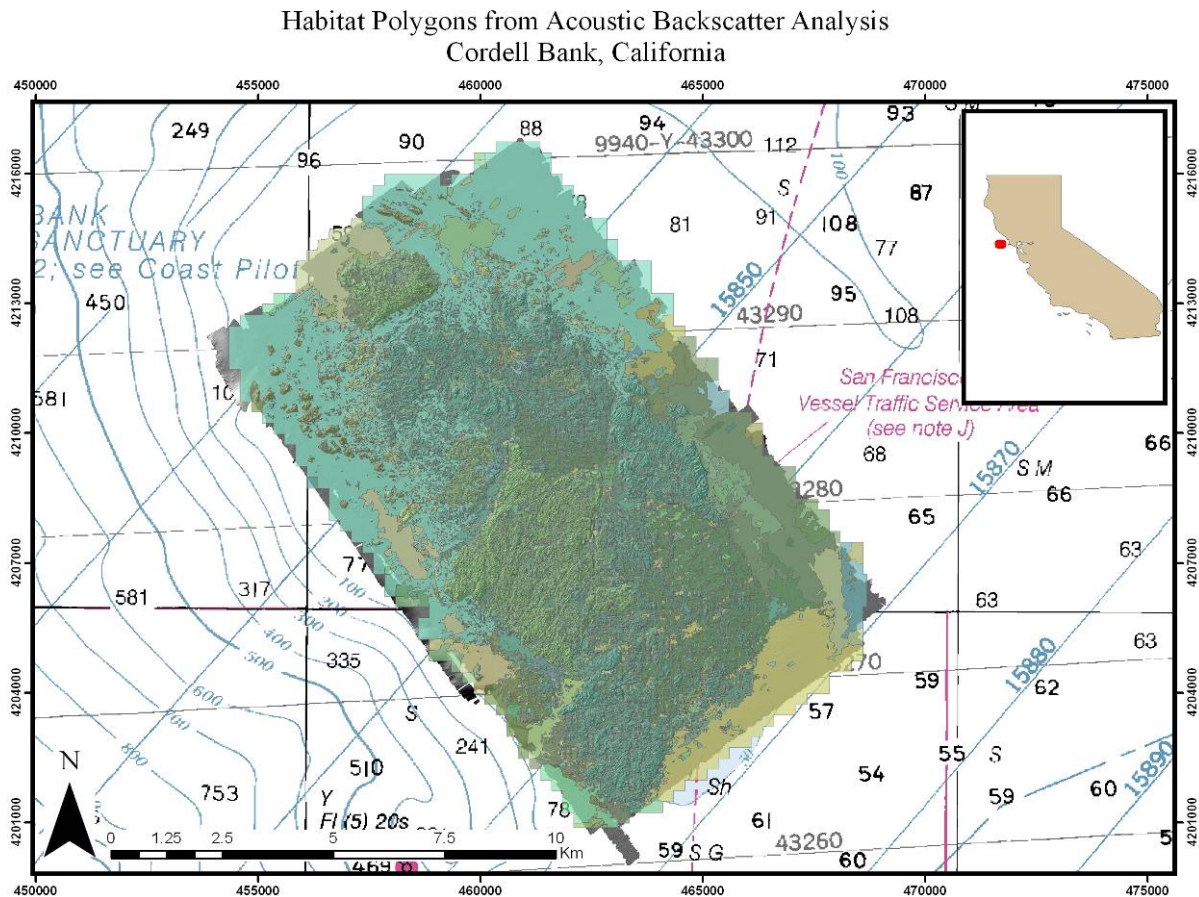
Figure 9: Habitat polygons generated from the acoustic backscatter analysis of Cordell Bank, California. The polygons are colored by the different habitat attributes. Over 100 habitat classes are included in this shapefile. (Coordinate System: UTM 10N WGS 84; NOAA Chart 18640, soundings in fathoms; survey completed fall 2005.)

The megahabitat categorization for Cordell Bank was defined as two different types. The area on the main bank was attributed with the letter "F" (flanks, continental slope, basin/island-atoll flanks), which comprises a total area of 80 km^2, while the area surrounding the bank was attributed with the letter "S" (shelf, continental and island shelves) and comprised a total area of 47 km^2 (Table 5).

Table 5: Megahabitat types for Cordell Bank, California from the Greene et al. deep water marine benthic habitat classification scheme (1999).

Megahabitat Type	Total Area (km^2)	Percentage of Site
F (flanks, continental slope, basin/island-atoll flanks)	80	63%
S (shelf, continental and island shelves)	47	37%

27

In general, the bank area consists of hard and mixed substrates while the area surrounding the bank consists of softer substrates with mixed substrate up toward the northern portion of the survey area (Figure 10).
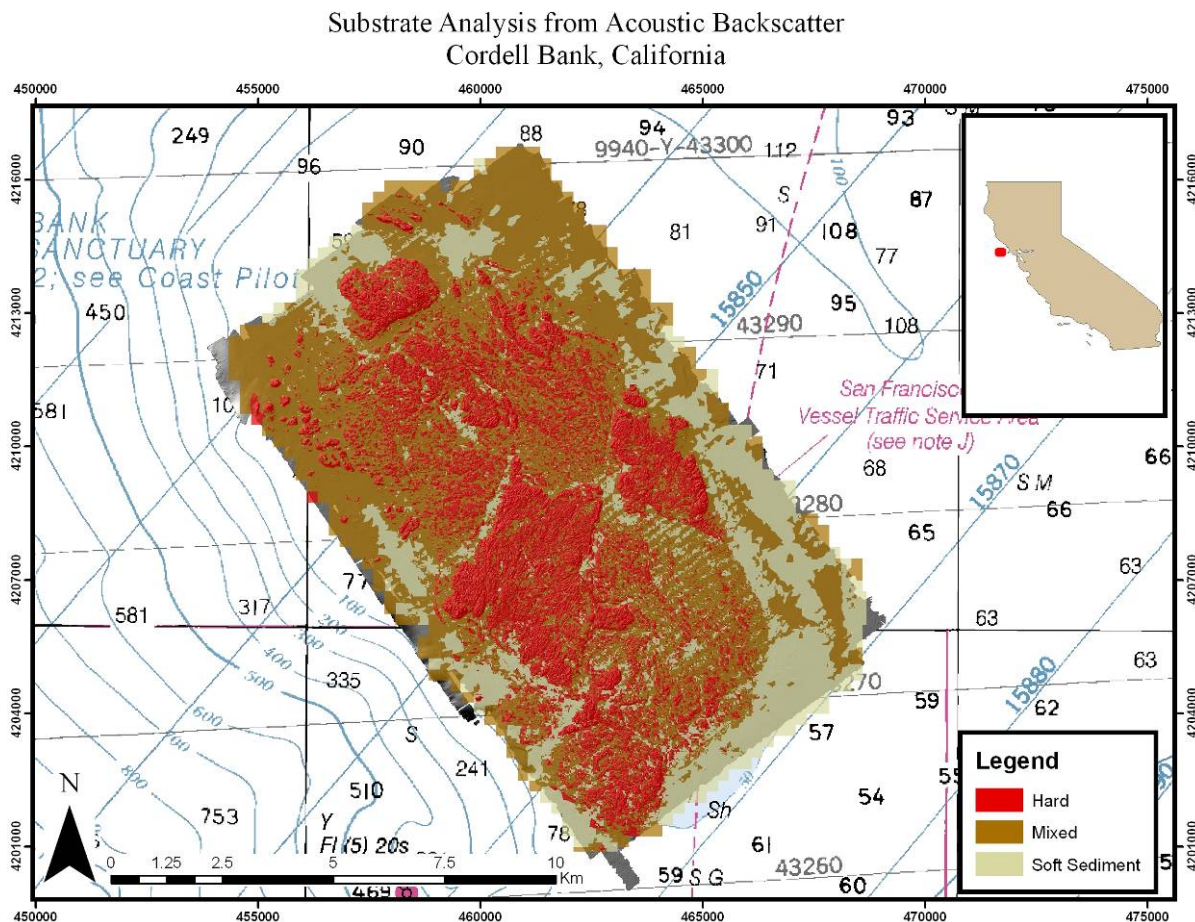
Figure 10: Substrate classification produced using the Maximum Likelihood Classification (MLC) performed on the acoustic backscatter form Cordell Bank, California. (Coordinate System: UTM 10N WGS 84; NOAA Chart 18640, soundings in fathoms; survey completed fall 2005.)

The different substrate types throughout Cordell Bank, California were given the following codes: "h" for hard substrates, "m" for mixed substrates, and "s" for soft substrates. Mixed substrate makes up the greatest area of the bank (48%) followed by hard substrate (29%) and soft substrate (23%). The total areas of hard, mixed, and soft substrates from the acoustic backscatter analysis were 36 km^2, 61 km^2, and 29 km^2, respectively (Table 6).

Table 6: The substrate classification results from the acoustic backscatter analysis and the corresponding areas of each substrate class on Cordell Bank, California.

Substrate Type	Area (km^2)	Percentage of Survey Area
Hard (h)	36	29%
Mixed (m)	61	48%
Soft Sediment (s)	29	23%

3.5 *Rockfish/Habitat Associations Analysis Using GLMs*

(The complete results for only the most extreme cases (best and worst) of the models for each species of fish are included in this results section. The full results for all species in each block that were not included in this section can be found in the Appendix.)

Sebastes. flavidus (Yellowtail Rockfish)

The model results for *S. flavidus* varied from block to block. The average adjusted D^2 value for all blocks was 0.26. This is a fairly low average meaning that these models don't have very high levels of performance. However, even though these models had low D^2 values, they correctly predicted a relatively high percentage of the probability of occurrence for the evaluation points. The average percent correctly predicted for *S. flavidus* across all the blocks was 76%. Overall, more errors of commission took place than errors of omission. This means that the models were more likely to make an error where they predicted a high probability of finding *S. flavidus* in an area where no fish were observed during the video transects than they were to predict a low probability of finding *S. flavidus* in an area where fish were actually observed. Therefore, on average, the models slightly over-predict the probability of occurrence for *Sebastes flavidus* (Table 7).

Table 7: Summary of the accuracy of the generalized linear models (GLMs) for *S. flavidus*.

Block	Adjusted D^2	% Correctly Predicted	% Errors of Omission	% Errors of Commission
Area 1	0.34	85	4	11
Area 2	0.32	80	3	17
Area 3	0.08	44	45	12
Area 4	0.19	72	5	23
Area 5	0.26	76	12	12
Area 6	0.37	97	0	3
Mean ± sd	**0.26 ± 0.11**	**76 ± 17.8**	**11.5 ± 16.9**	**13 ± 6.7**

For S. flavidus, Area 6 provided the most accurate predictions. The adjusted D^2 value for this model is 0.37. In addition, it accurately predicted the probability of occurrence for 97% of the test points. No errors of omission occurred and only 3% of the absence test points fell into areas of high probability and caused errors of commission (Table 8).

Table 8: Contingency table of the data used to determine the probability accuracy from the GLM performed in Area 6 to predict the probability of finding *S. flavidus* throughout the block (n = 76).

Point Type	# of Points where model indicates		Total
	Low Probability	High Probability	
Present	0	15	15
Absent	59	2	61
Total	59	17	76

Fine scale slope position and bathymetry were found to be significant in the GLM used to predict the probability of occurrence of *S. flavidus* in Area 6. Bathymetry had a greater effect on the model with a p-value of 0.000 but fine scale slope position was still significant with a p-value of 0.022 (Table 9).

Table 9: Summary of the probability results from the GLM performed on Area 6 to predict the probability of finding *S. flavidus* throughout the block.

Coefficients:	Estimate	Std. Error	z-value	Pr(>|z|)	
(Intercept)	7.254	3.081	2.354	0.019	*
FS Slope Position	0.642	0.280	2.294	0.022	*
Bathymetry	0.135	0.038	3.527	0.000	***

From the histograms outputted for both variables used in the model, it is clear why bathymetry was found to be significant. In the bathymetry distribution histogram, most of the presence points were found in depths between 65 and 85m with large spikes in density occurring at 80m and 70m. On the other hand, the histogram for the fine scale slope position variable shows that there is a fairly even distribution of fish density within all the categories while most of the absent points are found in the slope position values between 0 and 2 (Figure 11).

Figure 11: Histograms for each of the variables used in the GLM to predict the probability of occurrence of *S. flavidus* in Area 6. The densities of the presence and absence of *S. flavidus* are compared across the values of the predictor rasters.

The GLM formula for *S. flavidus* in Area 6 is:

$$7.254 + 0.642(\text{FS Slope Position}) + 0.135(\text{Bathymetry})$$

This GLM predicts that a high percentage (79%) of the area within Area 6 is unlikely to contain *S. flavidus*. It predicts that only 9% of the total block has a greater than 60% probability of containing *S. flavidus* (Table 10, Figure 12).

Table 10: Summary of the probability results from the GLM performed on Area 6 to predict the probability of finding *S. flavidus* throughout the block.

Probability of Occurrence	Area (m^2)	Area (km^2)	Percentage of Area 6
0.0-0.1	12836097	12.836	64%
0.11-0.2	987687	0.988	5%
0.21-0.3	807741	0.808	4%
0.31-0.4	1199250	1.199	6%
0.41-0.5	1379097	1.379	7%
0.51-0.6	937521	0.938	5%
0.61-0.7	718776	0.719	4%
0.71-0.8	398331	0.398	2%
0.81-0.9	419985	0.420	2%
0.91-1.0	187515	0.188	1%
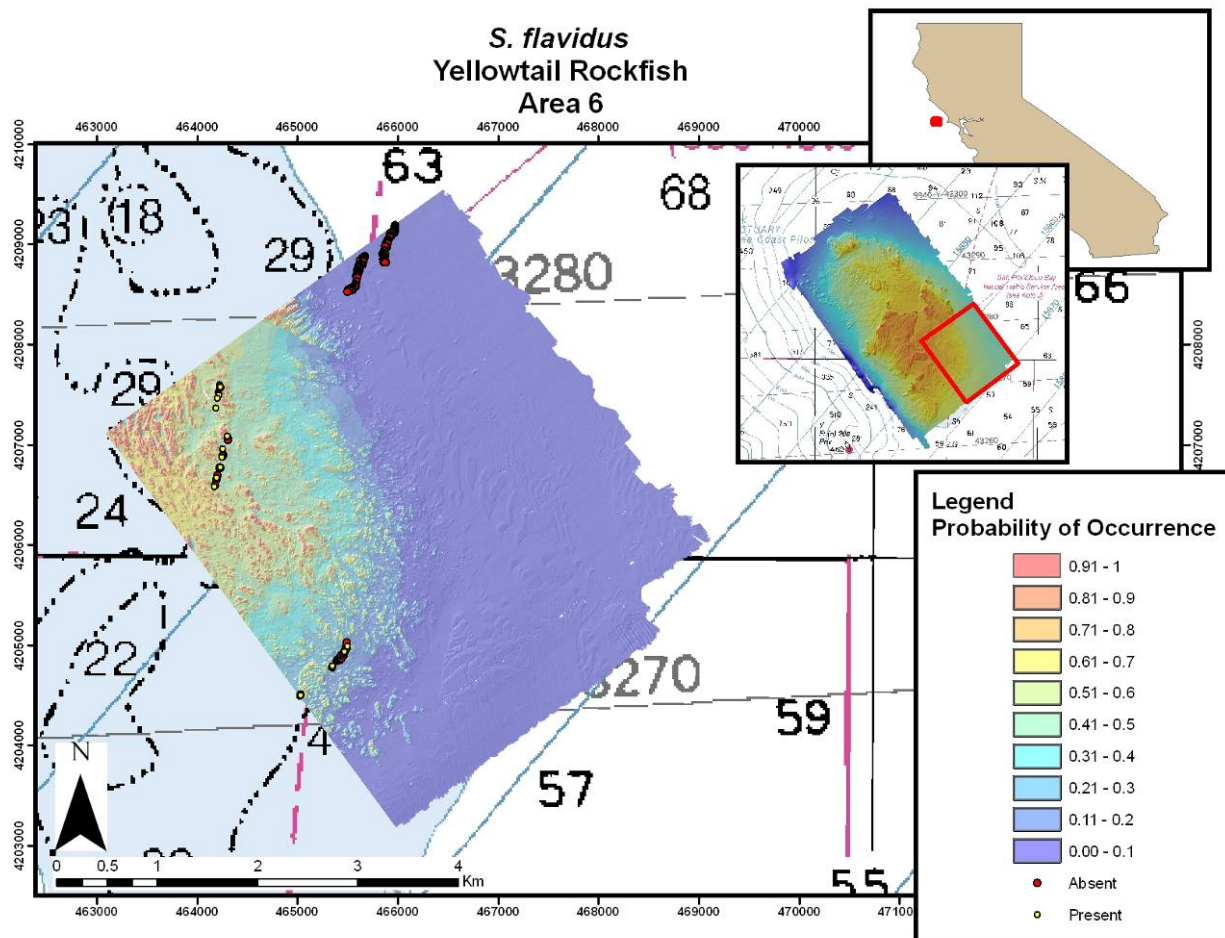
Figure 12: Area 6 GLM results for *S. flavidus* (yellowtail rockfish). Warmer colors indicate a high predicted probability of *S. flavidus* occurrence, while cooler colors indicate low probabilities. Yellow dots signify locations where *S. flavidus* were observed and the red dots indicate locations where no fish were present. (Image resolution: 3m; coordinate System: UTM 10N WGS 84; NOAA Chart 18640, soundings in fathoms; multibeam survey completed fall 2005.)

Unlike Area 6, the GLM for Area 3 did not accurately predict the presence/absence of a very high percentage of the evaluation points. This model has a very low adjusted D^2 value of 0.08. In addition, it only accurately predicted the probability of occurrence for 44% of the evaluation points. Errors of omission were made 45% of the time while errors of commission were made 12% of the time. Therefore, this model tended to predict low probability for areas where *S. flavidus* were observed during the video transects (Table 11).

Table 11: Contingency table of the data used to determine the probability accuracy from the GLM performed in Area 6 to predict the probability of finding *S. flavidus* throughout the block (n = 271).

| Point Type | # of Points where model indicates | | Total |
	Low Probability	High Probability	
Present	121	67	188
Absent	51	32	83
Total	172	99	271

Bathymetry, fine scale slope position, and aspect were the three variables found significant in the GLM for *S. flavidus* in Area 3. Bathymetry was the most significant with a p-value of <0.000 followed by the fine scale slope position (p=0.026) and then aspect (p=0.037) (Table 12).

Table 12: Summary of the predictor variables used in the GLM performed on Area 3 to predict the probability of finding *S. flavidus* throughout the block.

Coefficients:	Estimate	Std. Error	z value	Pr(>|z|)	
(Intercept)	3.230	0.807	4.000	6.33E-5	***
Bathymetry	0.034	0.006	5.597	2.18E-8	***
FS Slope Position	-0.226	0.102	-2.226	0.026	*
Aspect	0.003	0.001	2.083	0.037	*

The low levels of accuracy for this model can be explained by the histograms outputted for each of the variables. All of the presence points for both bathymetry and aspect had a fairly even distribution with points falling into every value of each variable. Although there were spikes of high density at certain values within the fine scale slope position raster, these spikes do not appear to be significantly different than those that occurred with the absence points (Figure 13).
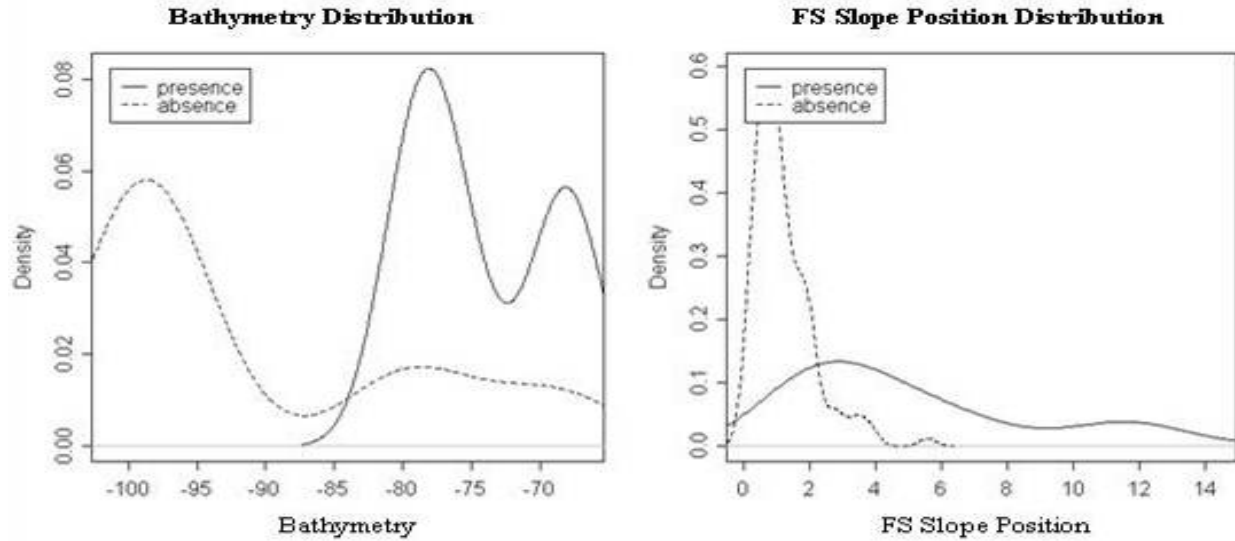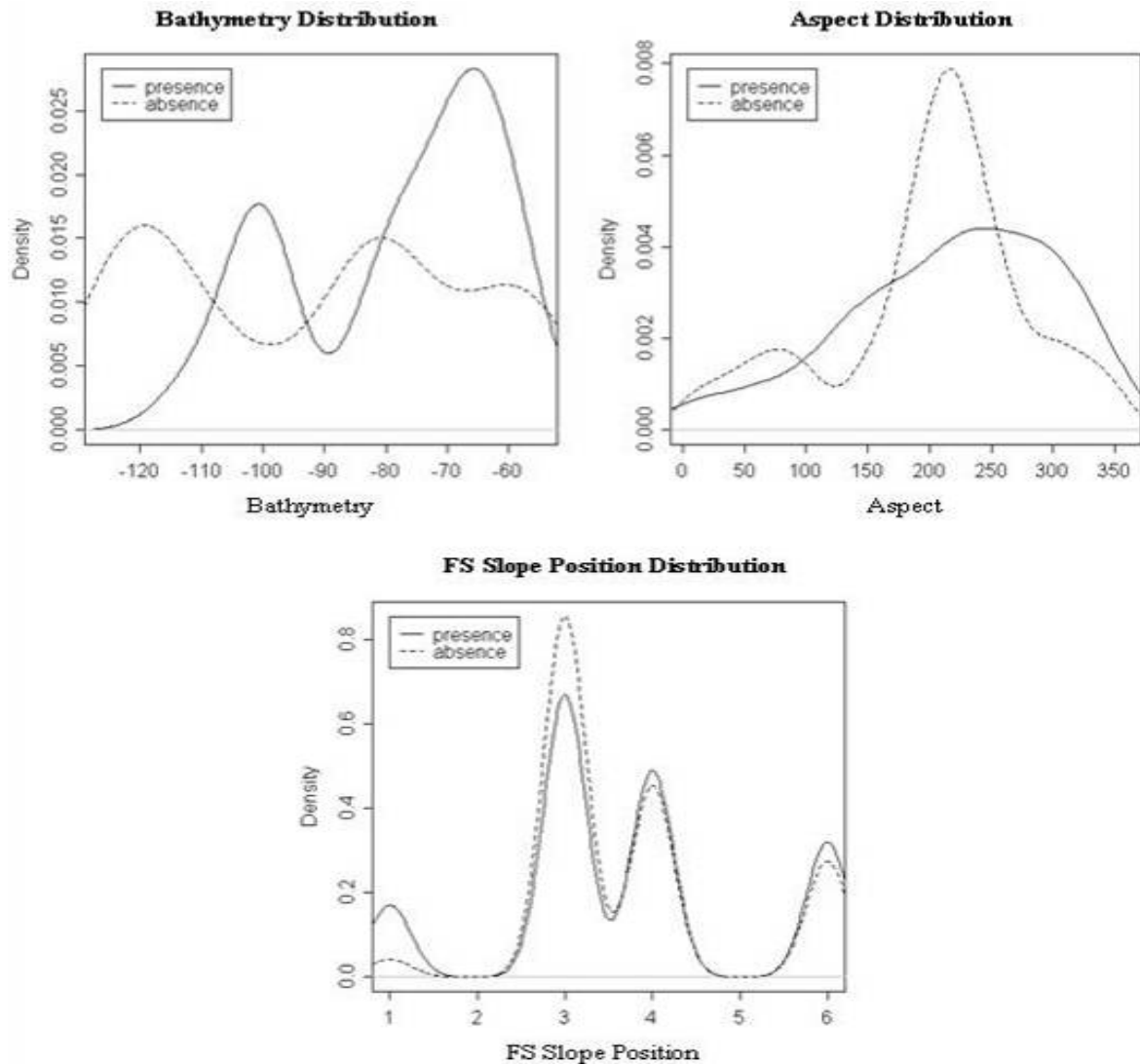
Figure 13: Histograms for each of the variables used in the GLM to predict the probability of occurrence of *S. flavidus* in Area 3. The densities of the presence and absence of *S. flavidus* are compared across the values of the predictor rasters.

The GLM formula for *S. flavidus* in Area 3 is:

$$3.230 + 0.034(\text{Bathymetry}) - 0.226(\text{FS Slope Position}) + 0.003 \,(\text{Aspect})$$

The percentage of Area 3 that fell into each of the probability ranges was fairly similar. The GLM only predicted that 31% of the block had a greater than 60% chance of containing *S. flavidus* and 38% were unlikely to contain *S. flavidus* (Table 13, Figure 14).

Table 13: Summary of the probability results from the GLM performed on Area 3 to predict the probability of finding *Sebastes flavidus* throughout the block.

| Probability of Occurrence	Area (m^2)	Area (km^2)	Percentage of Area 3
0.0-0.1	779193	0.779	3%
0.11-0.2	1344339	1.344	6%
0.21-0.3	3716712	3.717	16%
0.31-0.4	3462363	3.462	15%
0.41-0.5	3391866	3.392	15%
0.51-0.6	3610170	3.610	16%
0.61-0.7	3641526	3.642	16%
0.71-0.8	2726550	2.727	12%
0.81-0.9	585135	0.585	3%
0.91-1.0	855	0.001	0%
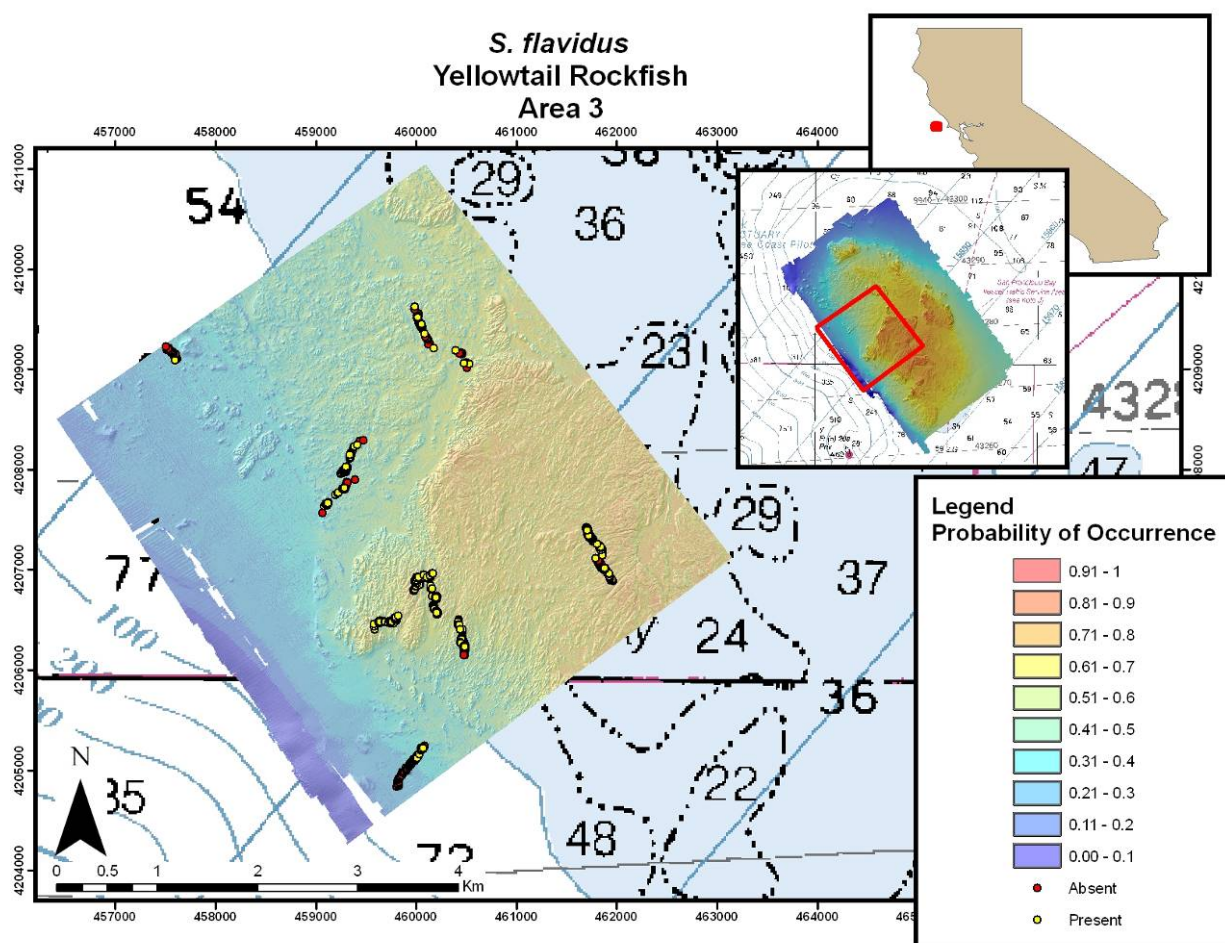
Figure 14: Area 3 GLM results for *S. flavidus* (yellowtail rockfish). Warmer colors indicate a high predicted probability of *S. flavidus* occurrence, while cooler colors indicate low probabilities. Yellow dots signify locations where *S. flavidus* were observed and the red dots indicate locations where no fish were present. (Image resolution: 3m; coordinate System: UTM 10N WGS 84; NOAA Chart 18640, soundings in fathoms; multibeam survey completed fall 2005.)

When all of the models were combined together for *S. flavidus* across the entire site, it is obvious that different predictions were made in each of the blocks and a seamless model cannot be achieved for all of Cordell Bank. Even in areas where two blocks slightly overlap, there are distinguishable differences in the types of habitat that are found to be the most suitable. This is most obvious between Areas 1 and 3. The majority of Area 1 contains habitat that has a >90% probability of occurrence for S. flavidus while Area 3 only contains a very insignificant amount of habitat that has a probability of occurrence between 0.91 and 1.0. When the probability rasters for these two areas are placed next to each other, there is an observable seam. The rest of the areas have models that are slightly different but the differences are less apparent (Figure 15).
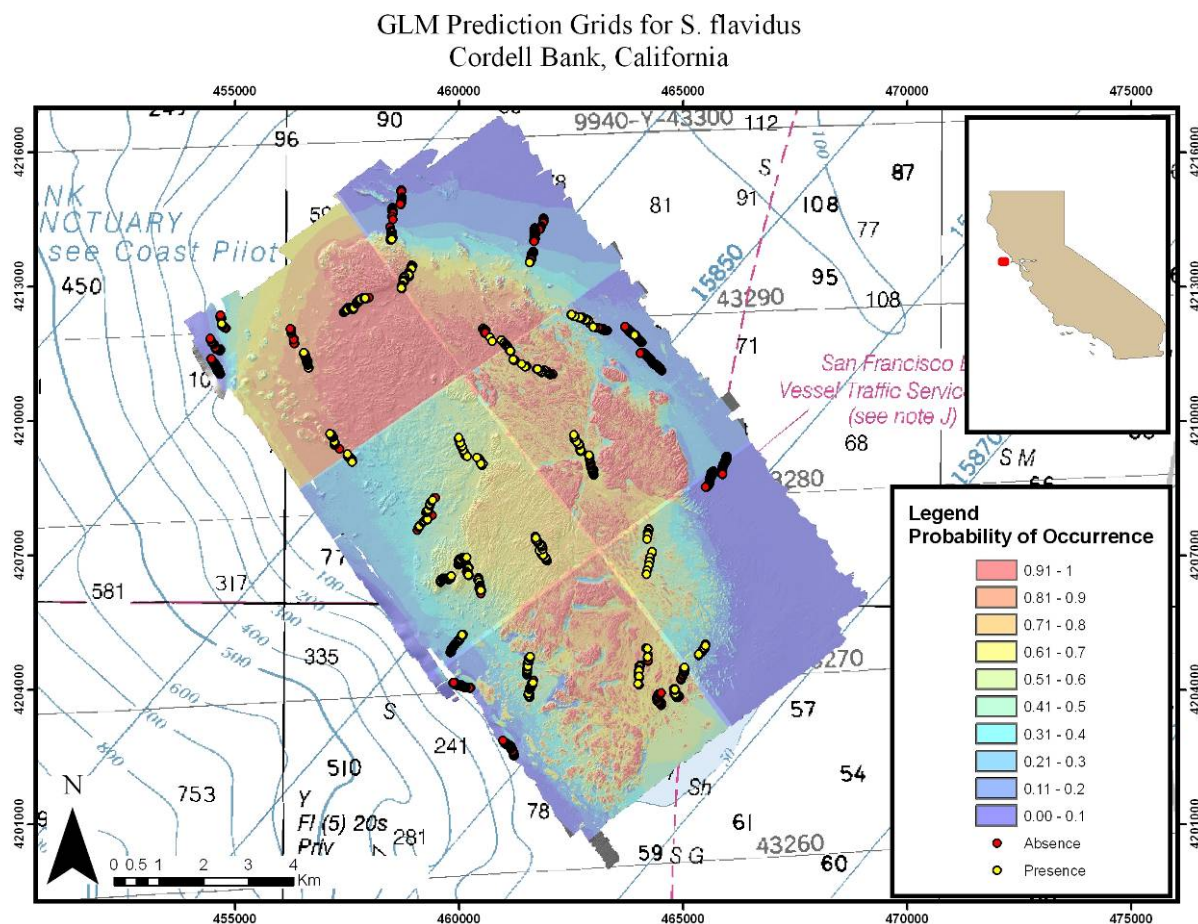
Figure 15: A combination of the probability of occurrence rasters for *S. flavidus* across the entire site. Warmer colors indicate a high predicted probability of *S. flavidus* occurrence, while cooler colors indicate low probabilities. Yellow dots signify locations where *S. flavidus* were observed and the red dots indicate locations where no fish were present. (Image resolution: 3m; coordinate System: UTM 10N WGS 84; NOAA Chart 18640, soundings in fathoms; multibeam survey completed fall 2005.)

36

Sebastes rosaceus (Rosy Rockfish)

The GLMs that were ran on the presence and absence points for *S. rosaceus* had relatively high levels of performance. The average adjusted D^2 value for these models was 0.45. The highest level of performance was in Area 4 with an adjusted D^2 value equal to 0.66. The generalized linear model for Area 3 had the lowest level of performance (adjusted $D^2 = 0.08$).

Along with having high levels of performance, the GLMs for *S. rosaceus* proved to be fairly accurate. They correctly predicted the probability of occurrence 81% of the time on average. Area 2 and Area 6 had the highest levels of accuracy with 91% and 92% correct predictions, respectively. Area 3 had the lowest level of accuracy with only 63% of the test points being predicted correctly. Errors of omission and errors of commission occurred about the same percentage of the time with errors of omission occurring slightly more on average. Therefore, the models for *S. rosaceus* tend to predict low levels of probability where fish are actually present. This case is most apparent in area three where 25% of the presence points from the test transects fell in areas with low probabilities of occurrence while only 12% of the absence points fell in areas with high probabilities of occurrence (Table 14).

Table 14: Summary of the accuracy of the generalized linear models (GLMs) for *S. rosaceus*.

Block	Adjusted D^2	% Correctly Predicted	% Errors of Omission	% Errors of Commission
Area 1	0.52	79	5	16
Area 2	0.55	91	7	2
Area 3	0.08	63	25	12
Area 4	0.66	80	7	13
Area 5	0.38	81	9	10
Area 6	0.53	92	7	1
Mean ± sd	**0.45 ± 0.20**	**81 ± 10.5**	**10 ± 7.5**	**9 ± 6.1**

Although most of the GLMs for *S. rosaceus* were fairly accurate in predicting distribution, Area 2 was one of the most accurate with a high adjusted D^2 value of 0.55 and 91% of the probability of occurrence for the evaluation points being predicted correctly. Errors of omission were made 7% of the time and errors of commission were made 2% of the time (Table 15).

Table 15: Contingency table of the data used to determine the probability accuracy from the GLM performed in Area 2 to predict the probability of finding *Sebastes rosaceus* throughout the block (n = 184).

| Point Type | # of Points where model indicates | | Total |
	Low Probability	High Probability	
Present	12	21	33
Absent	146	5	151
Total	158	26	184

Only two variables were found to be significant in the GLM used to predict the probability of occurrence of *S. rosaceus* in Area 2. Bathymetry had the greatest effect with a p-value of 0.006 but aspect was still significant with a p-value of 0.026 (Table 16).

Table 16: Summary of the predictor variables used in the GLM performed on Area 2 to predict the probability of finding *S. rosaceus* throughout the block.

Coefficients:	Estimate	Std. Error	z-value	Pr(>|z|)	
(Intercept)	17.815	6.039	2.950	0.003	**
Bathymetry	0.201	0.073	2.768	0.006	**
Aspect	-0.006	0.003	-2.228	0.026	*

The bathymetry distribution histogram shows that *S. rosaceus* are found in depths between 75 and 95m with the greatest densities occurring at 80m. In addition, they are most commonly found on habitat that has an aspect around 50˚ (Figure 16).
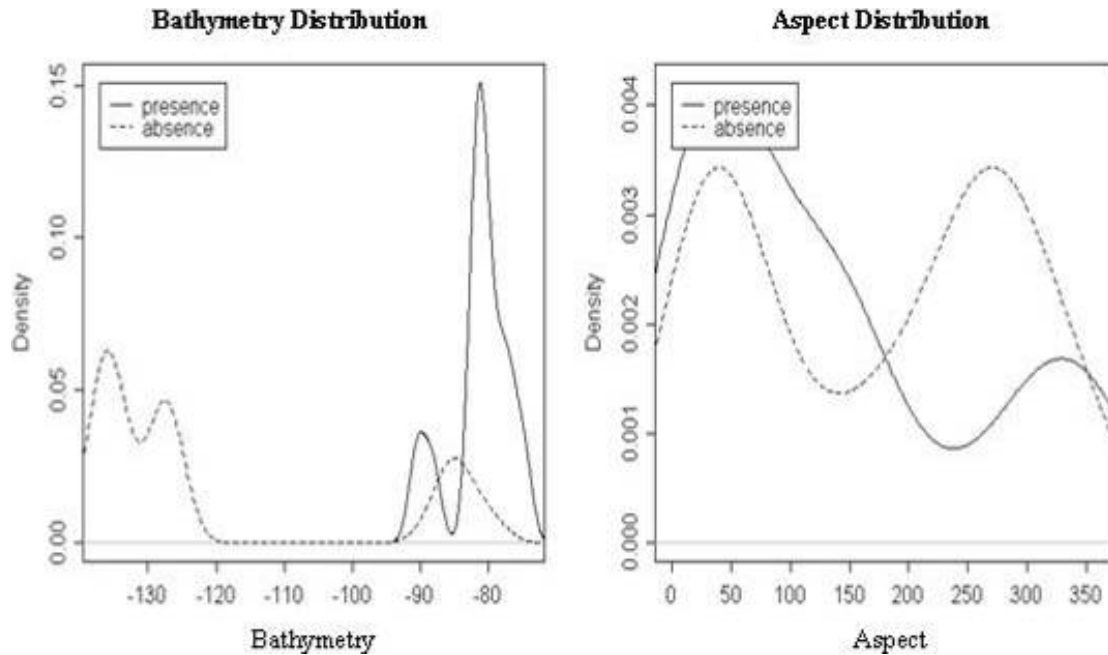
Figure 16: Histograms for each of the variables used in the GLM for Area 2. The densities of the presence and absence of *S. rosaceus* are compared across the values of the predictor rasters.

The GLM formula for *S. rosaceus* in Area 2 is:

$$17.815 + 0.201(\text{Bathymetry}) - 0.006(\text{Aspect})$$

The majority of the output from this GLM fell into the 0.0 – 0.1 range of probabilities (80%). Only 8% of the habitat in Area 2 had a high probability of occurrence for *S. rosaceus* (Table 17, Figure 17).

Table 17: Summary of the probability results from the GLM performed on Area 2 to predict the probability of finding *S. rosaceus* throughout the block.

Probability of Occurrence	Area (m^2)	Area (km^2)	Percentage of Area 2
0.0-0.1	15361218	15.36122	80%
0.11-0.2	671490	0.67149	3%
0.21-0.3	445338	0.445338	2%
0.31-0.4	362898	0.362898	2%
0.41-0.5	357768	0.357768	2%
0.51-0.6	361683	0.361683	2%
0.61-0.7	378315	0.378315	2%
0.71-0.8	420390	0.42039	2%
0.81-0.9	440073	0.440073	2%
0.91-1.0	398349	0.398349	2%
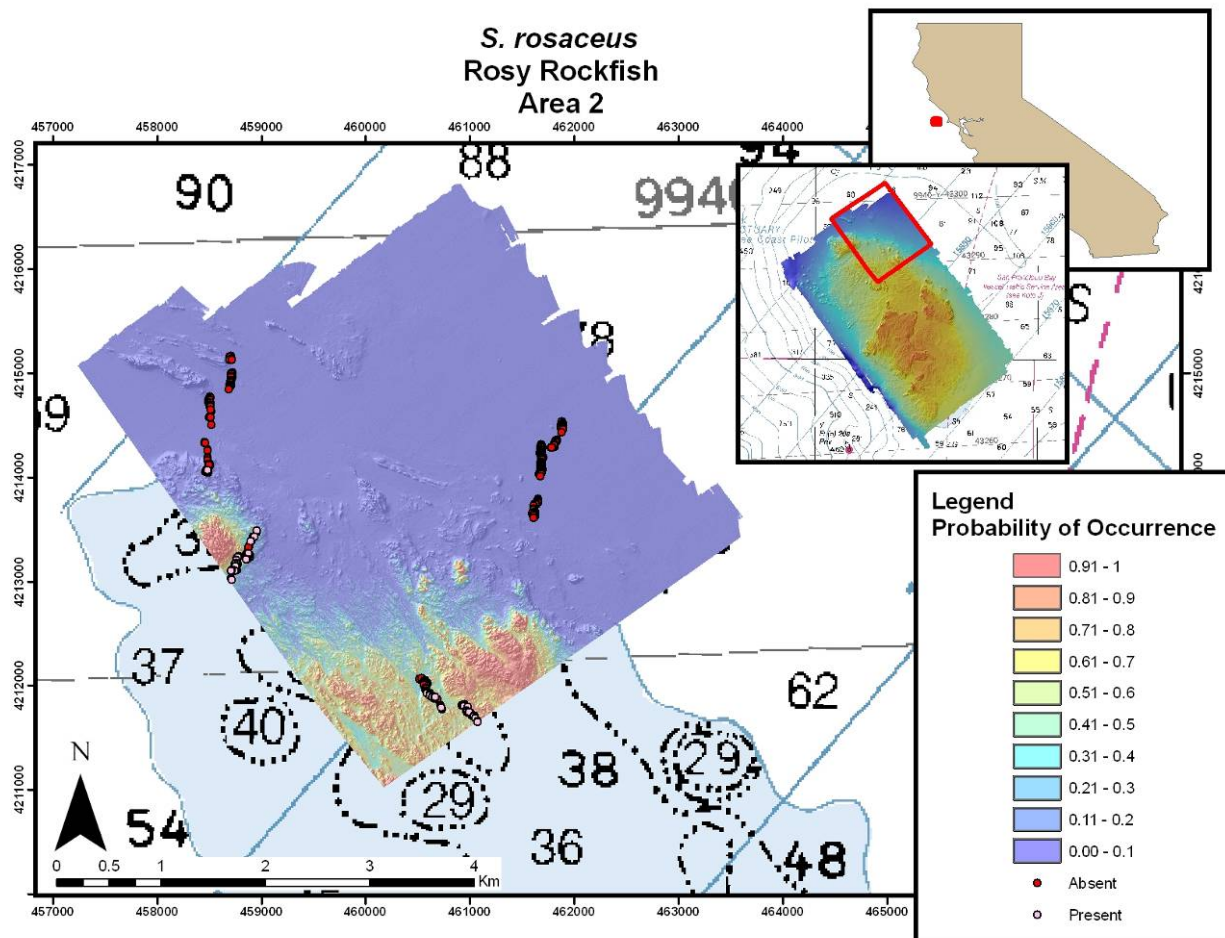
Figure 17: Area 2 GLM results for *S. rosaceus* (rosy rockfish). Warmer colors indicate a high predicted probability of *S. rosaceus* occurrence, while cooler colors indicate low probabilities. Pink dots signify locations where *S. rosaceus* were observed and the red dots indicate locations where no fish were present. (Image resolution: 3m; coordinate System: UTM 10N WGS 84; NOAA Chart 18640, soundings in fathoms; multibeam survey completed fall 2005.)

As with *S. flavidus*, the GLM performed in Area 3 for *S. rosaceus* came up with poor results. This model has an adjusted D^2 value of 0.08 meaning that 92% of the residual deviance from this model was not explained by the model. It accurately predicted the probability of occurrence for *S. rosaceus* in Area 3 63% of the time. Errors of omission occurred for 25% of the evaluation points and errors or commission occurred for 12% of the evaluation points. Therefore, this model tended to predict low probabilities of occurrence in areas where *S. rosaceus* were observed during the video transects (Table 18).

Table 18: Contingency table of the data used to determine the probability accuracy from the GLM performed in Area 2 to predict the probability of finding *S. rosaceus* throughout the block (n = 179).

Point Type	# of Points where model indicates		Total
	Low Probability	High Probability	
Present	45	57	102
Absent	56	21	77
Total	101	78	179

The generalized linear model used to predict the probability of occurrence for *S. rosaceus* in Area 3, only found bathymetry and aspect to be significant with p-values of 0.000 and 0.042, respectively (Table 19).

Table 19: Summary of the predictor variables used in the GLM performed on Area 3 to predict the probability of finding *S. rosaceus* throughout the block.

Coefficients:	Estimate	Std. Error	z value	Pr(>|z|)	
(Intercept)	2.094	0.746	2.806	0.005	**
Aspect	0.003	0.002	2.037	0.042	*
Bathymetry	0.033	0.009	3.770	0.000	***

The distributions of the presence points for *S. rosaceus* are fairly even over the values from both predictor rasters. There are no significant spikes in density for any of the values from either variable (Figure 18).
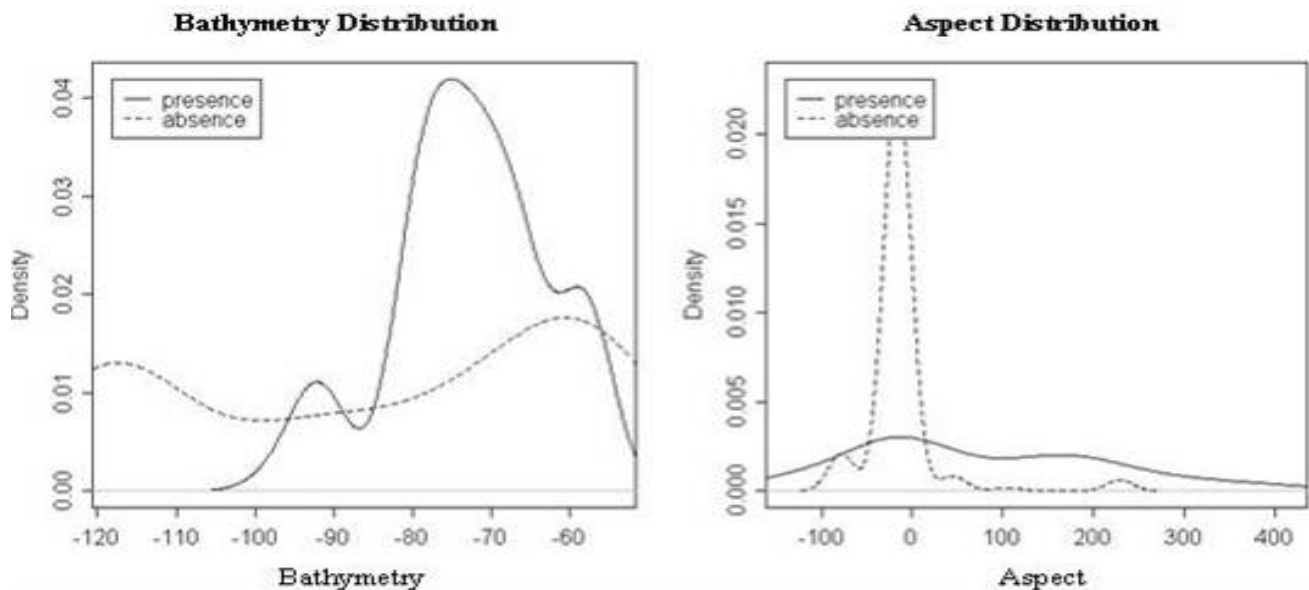
Figure 18: Histograms for each of the variables used in the GLM to predict the probability of occurrence of *S. rosaceus* in Area 2. The densities of the presence and absence of *S. rosaceus* are compared across the values of the predictor rasters.

The GLM formula for *S. rosaceus* in Area 3 is:

$$2.094 + 0.003(\text{Aspect}) + 0.033(\text{Bathymetry})$$

The results from the GLM indicate that 28% of Area 3 has a high probability of containing *S. rosaceus* and that 43% of the site has a small probability of being suitable habitat for *S. rosaceus* (Table 20, Figure 19).

Table 20: Summary of the probability results from the GLM performed on Area 3 to predict the probability of finding *S. rosaceus* throughout the block.

| Probability of Occurrence	Area (m^2)	Area (km^2)	Percentage of Area 3
0.0-0.1	695043	0.695	3%
0.11-0.2	2026107	2.026	9%
0.21-0.3	3816450	3.816	17%
0.31-0.4	3272733	3.272	14%
0.41-0.5	3223674	3.223	14%
0.51-0.6	3634371	3.634	16%
0.61-0.7	3533031	3.533	15%
0.71-0.8	2814831	2.814	12%
0.81-0.9	242469	0.242	1%
0.91-1.0	0	0	0%
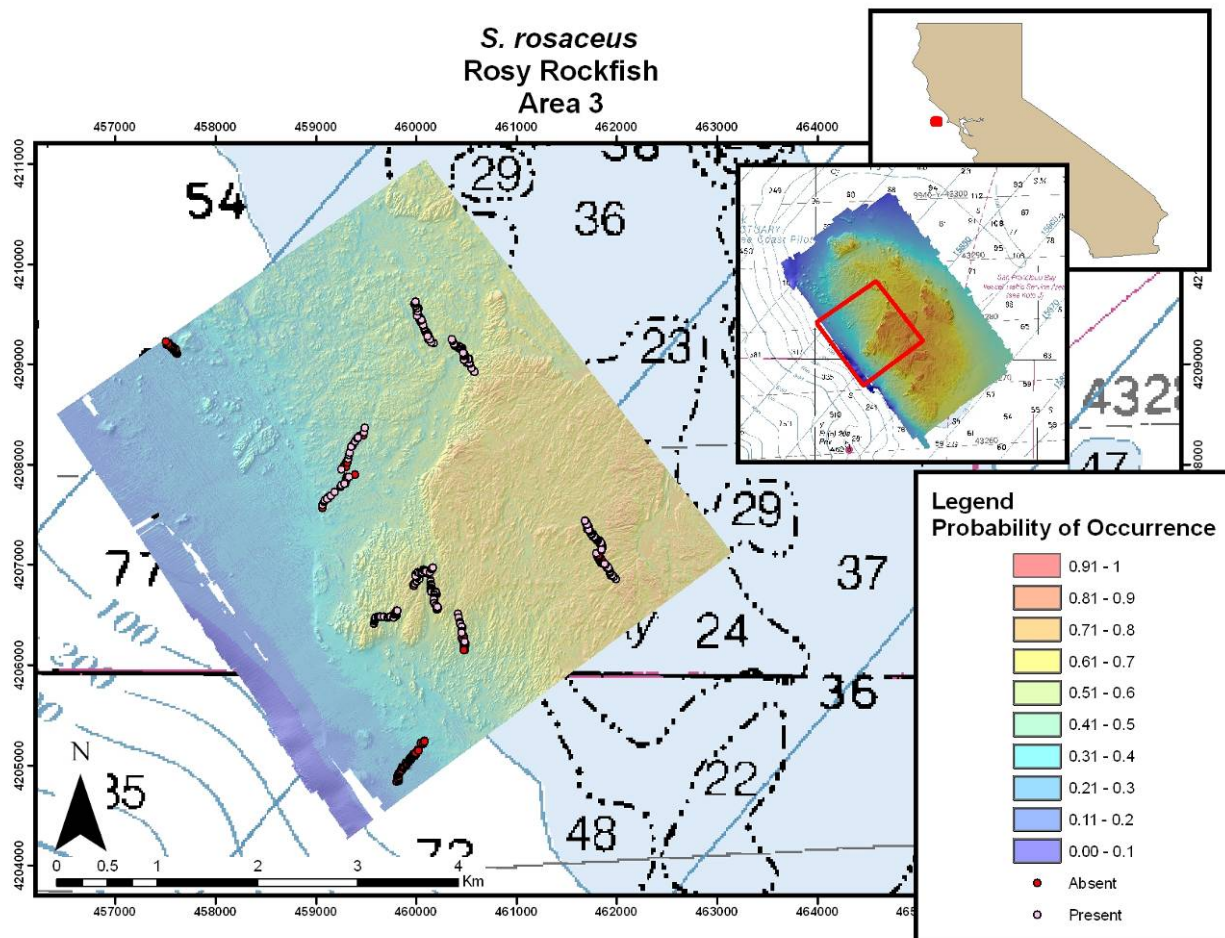
Figure 19: Area 3 GLM results for *S. rosaceus* (rosy rockfish). Warmer colors indicate a high predicted probability of *S. rosaceus* occurrence, while cooler colors indicate low probabilities. Pink dots signify locations where *S. rosaceus* were observed and the red dots indicate locations where no fish were present. (Image resolution: 3m; coordinate System: UTM 10N WGS 84; NOAA Chart 18640, soundings in fathoms; multibeam survey completed fall 2005.)

When all of the GLM prediction grids are combined across the entire site, there are some obvious discrepancies between the predictions in the separate blocks but the majority of the areas seem to match up fairly well. The main discrepancies occur between areas 1 and 3. There are obvious seams where these two areas meet. However, the rest of the blocks seem to match up fairly well, especially areas 2, 4, and 6 on the east side of the bank. These three prediction grids can be combined to create an almost seamless prediction grid (Figure 20).

GLM Prediction Grids for S. rosaceus
Cordell Bank, California
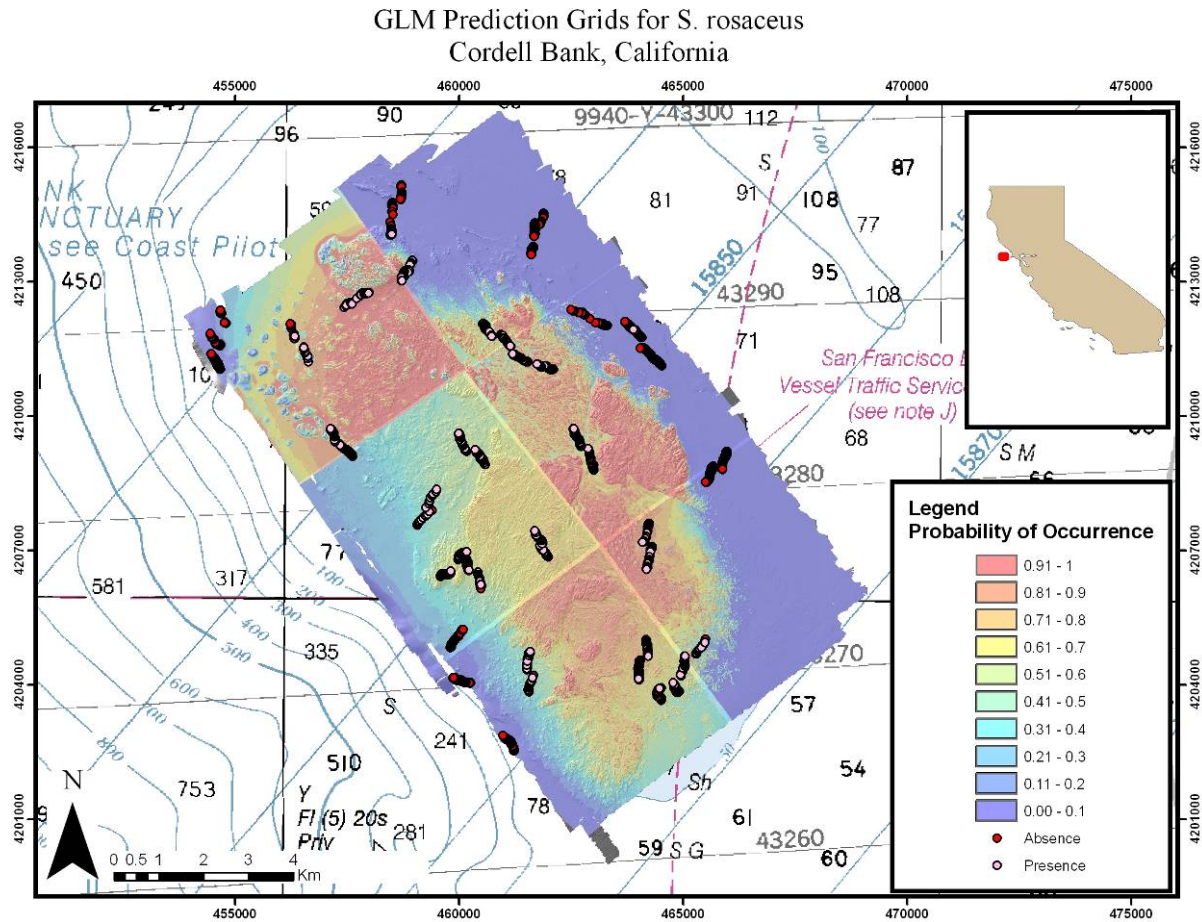
Figure 20: A combination of the probability of occurrence rasters for *Sebastes rosaceus* across the entire site. Warmer colors indicate a high predicted probability of *S. rosaceus* occurrence, while cooler colors indicate low probabilities. Pink dots signify locations where *S. rosaceus* were observed and the red dots indicate locations where no fish were present. (Image resolution: 3m; coordinate System: UTM 10N WGS 84; NOAA Chart 18640, soundings in fathoms; multibeam survey completed fall 2005.)

Sebastes elongatus (Greenstriped Rockfish)

The distribution of *S. elongatus* proved to be relatively difficult to model. The average adjusted D^2 value for this species was 0.25. This is a relatively low average and means that about 75% of the residual deviance could not be explained by these models. Therefore, the GLM for either of the areas where *S. elongatus* were present did not have very high levels of performance.

The average percentage of the test points that were correctly predicted was 62%. Only 45% of the points were correctly predicted in Area 1 while 78% were correctly predicted in Area 2. Each of these areas experienced different patterns in their errors of omission versus their errors of commission. In Area 1, 6% of the presence points fell into areas of low probability of occurrence while 48% of the absence points fell into areas of high probability of occurrence. On the other hand, in Area 2, 22% of the presence points fell into areas of low probability of occurrence and none of the absence points fell in areas of high probability of occurrence. These results show that the GLM for *S. elongatus* in Area 1 tended to over-predict the probability of occurrence while the GLM for *S. elongatus* in Area 2 tended to under-predict the probability of occurrence (Table 21).

Table 21: Summary of the accuracy of the generalized linear models (GLMs) for *S. elongatus*.

Block	Adjusted D^2	% Correctly Predicted	% Errors of Omission	% Errors of Commission
Area 1	0.25	45	6	48
Area 2	0.25	78	22	0
Mean ± sd	**0.25 ± 0**	**62 ± 23.3**	**14 ± 11.3**	**24 ± 33.9**

Neither GLM for *S. elongatus* was very effective at predicting the probability of occurrence for this species but the model for Area 1 was slightly better than that for Area 2. The adjusted D^2 value calculated for this GLM is 0.25. This model was only 45% accurate in predicting the probability of occurrence of *S. elongatus*. Errors of omission occurred for 6% of the test points while 48% errors of commission occurred (Table 22).

Table 22: Contingency table of the data used to determine the probability accuracy from the GLM performed in Area 1 to predict the probability of finding *S. elongatus* throughout the block (n = 93).

Point Type	# of Points where model indicates		Total
	Low Probability	High Probability	
Present	6	6	12
Absent	36	45	81
Total	42	51	93

The GLM used to predict the probability of occurrence of *S. elongatus* within Area 1, found the following variables to be significant: rugosity (p=0.016), broad scale slope position (p=0.041), bathymetry (p=0.029), and substrate type from the acoustic backscatter maximum

likelihood classification (MLC; p=0.006; Table 23). This was the only model where the MLC of the acoustic backscatter was found to be a significant predictor.

Table 23: Summary of the predictor variables used in the GLM performed on Area 1 to predict the probability of finding *S. elongatus* throughout the block.

Coefficients:	Estimate	Std. Error	z value	Pr(>|z|)	
(Intercept)	69.858	43.224	1.616	0.016	**
Rugosity	-80.207	44.797	1.548	0.050	*
BS Slope Position	-0.962	0.471	-2.044	0.041	*
Bathymetry	-0.040	0.018	-2.179	0.029	*
MLC	1.942	0.701	2.772	0.006	**

The histograms from the GLM used to predict the probability of occurrence of *S. elongatus* in Area 1 display the distribution of the presence/absence points over the values of the predictor rasters. There are fairly even distributions within bathymetry and broad scale slope position. On the other hand, *S. elongatus* is found on a very narrow range of the values from the rugosity and the MLC. They were only present in the very low rugosity values and the majority of the fish were found on mixed substrate from the MLC (Figure 21).
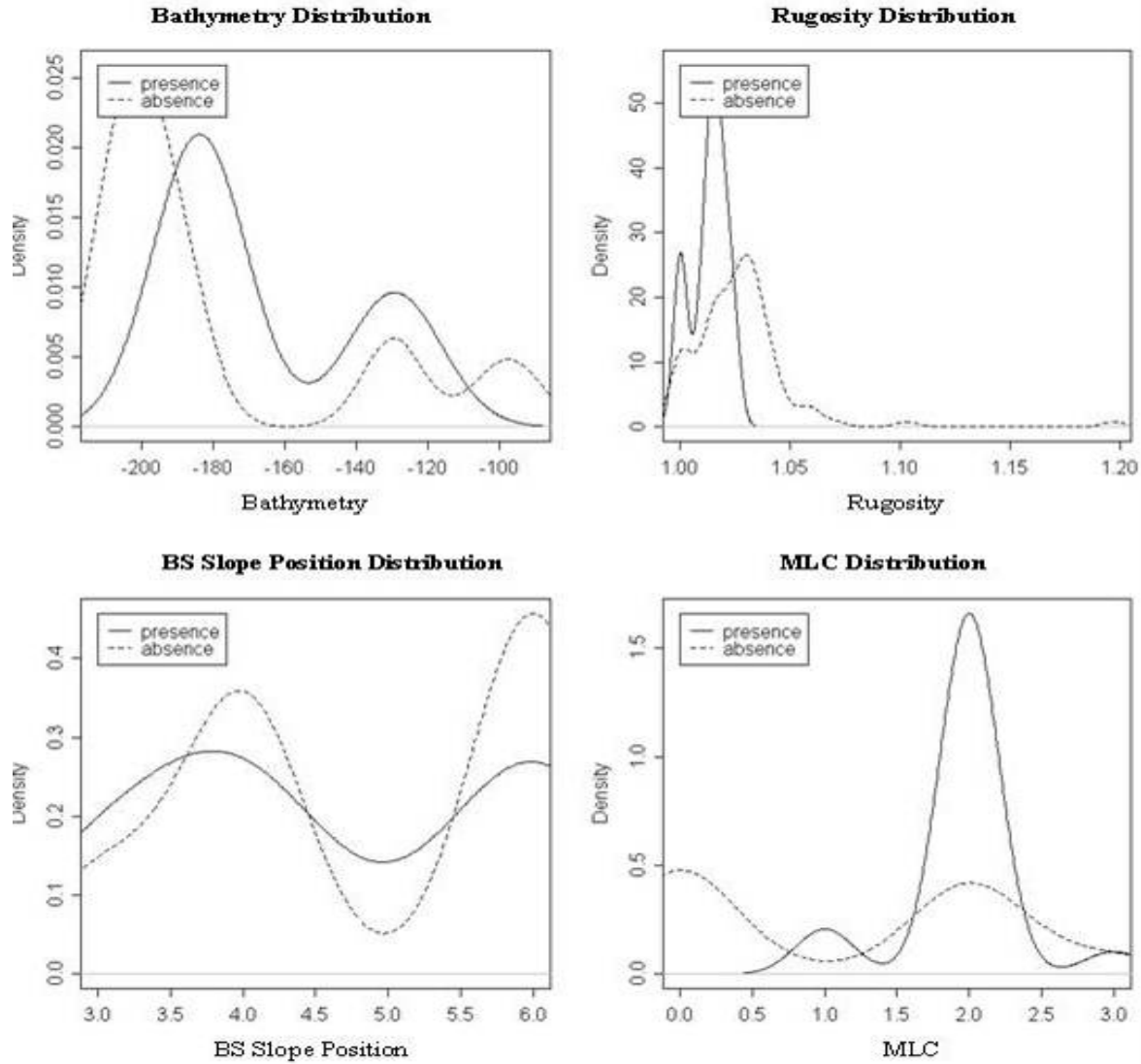
Figure 21: Histograms for each of the variables used in the GLM for Area 1. The densities of the presence and absence of *Sebastes elongatus* are compared across the values of the predictor rasters.

The GLM formula for *S. elongatus* in Area 3 is:

69.858 – 80.207(rugosity) – 0.962(BS Slope Position) – 0.040(Bathymetry) + 1.942(MLC)

This GLM predicted that 62% of the habitat within Area 1 had a less than 40% chance of containing *S. elongatus*. Only 15% of the area is likely to have *S. elongatus* present (Table 24, Figure 22).

Table 24: Summary of the probability results from the GLM performed on Area 1 to predict the probability of finding *S. elongatus* throughout the block.

Probability of Occurrence	Area (m²)	Area (km²)	Percentage of Area 1
0.0-0.1	3771036	3.771	19%
0.11-0.2	3586707	3.587	18%
0.21-0.3	2156652	2.157	11%
0.31-0.4	2745378	2.745	14%
0.41-0.5	2756871	2.757	14%
0.51-0.6	1643607	1.644	8%
0.61-0.7	1133703	1.134	6%
0.71-0.8	853605	0.854	4%
0.81-0.9	560754	0.561	3%
0.91-1.0	429156	0.429	2%
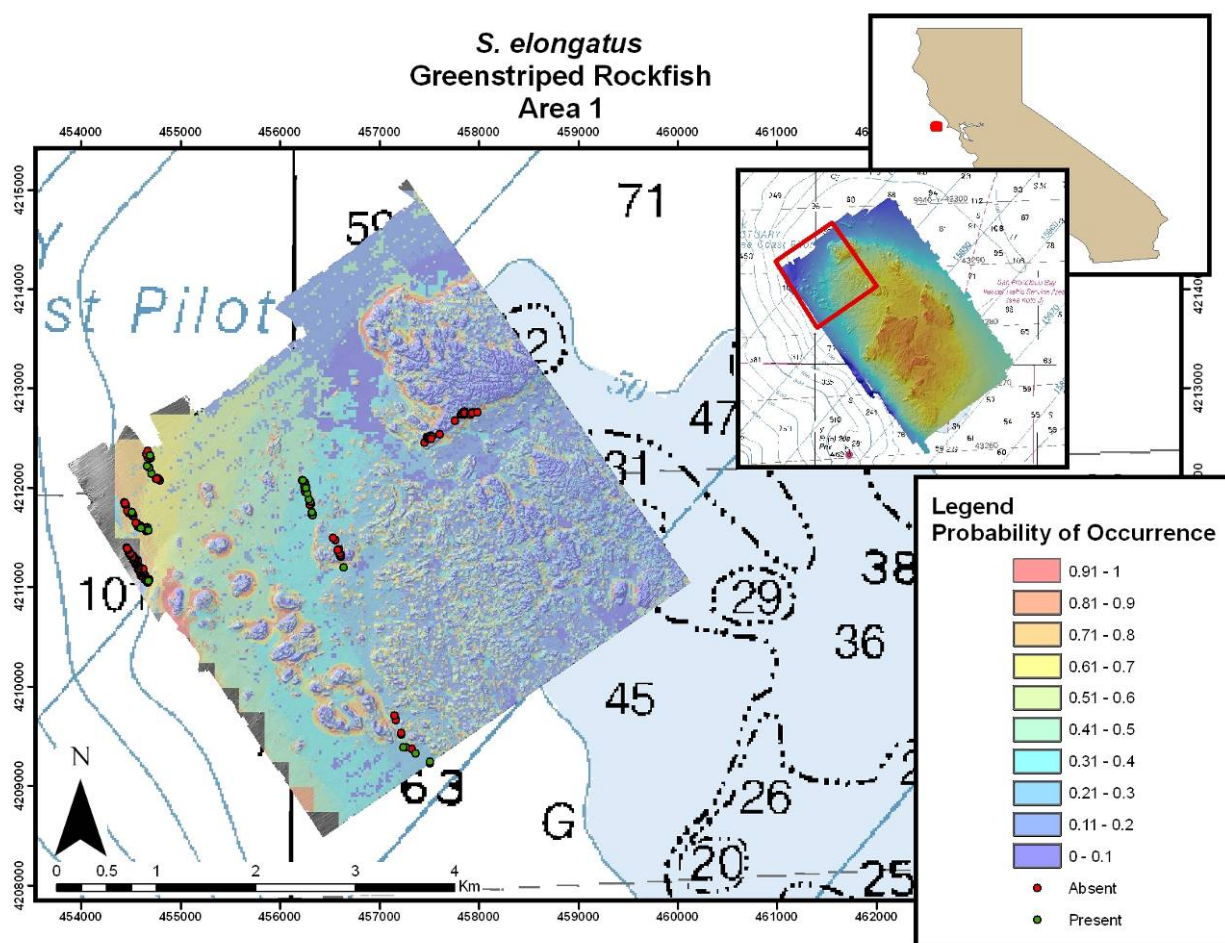
Figure 22: Area 1 GLM results for *S. elongatus* (greenstriped rockfish). Warmer colors indicate a high predicted probability of *S. elongatus* occurrence, while cooler colors indicate low probabilities. Green dots signify locations where *S. elongatus* were observed and the red dots indicate locations where no fish were present. (Image resolution: 3m; coordinate System: UTM 10N WGS 84; NOAA Chart 18640, soundings in fathoms; multibeam survey completed fall 2005.)

Although the GLM for *S. elongatus* in Area 1 was not very effective in predicting the probability of occurrence, the model for Area 2 was even less effective. The adjusted D^2 value calculated for this GLM is 0.25. None of the points for this test fell into an area that had a high probability of occurrence for *S. elongatus*. All of the test points for presence were found in an area of low probability, making the model have a 22% error of omission. No errors of commission occurred. This model had an overall accuracy of 78% (Table 25).

Table 25: Contingency table of the data used to determine the probability accuracy from the GLM performed in Area 2 to predict the probability of finding *S. elongatus* throughout the block (n = 166).

| Point Type | # of Points where model indicates | | Total |
	Low Probability	High Probability	
Present	37	0	37
Absent	129	0	129
Total	166	0	166

The following variables were used in the GLM to predict the probability of occurrence of *S. elongatus* in Area 2: bathymetry (p<0.000) and broad scale slope position (p=0.000) (Table 26).

Table 26: Summary of the predictor variables used in the GLM performed on Area 2 to predict the probability of finding *S. elongatus* throughout the block.

Coefficients:	Estimate	Std. Error	z value	Pr(>|z|)	
(Intercept)	-52.868	24.729	-2.138	0.033	*
Bathymetry	-0.353	0.079	-4.467	7.94E-6	***
BS Slope Position	-2.262	0.619	-3.653	0.000	**

The bathymetry distribution histogram from this GLM has its largest spike in density of *S. elongatus* at 140m of depth. However, the broad scale slope position distribution shows that there is an almost even distribution in density of *S. elongatus* across all values of this predictor (Figure 23).
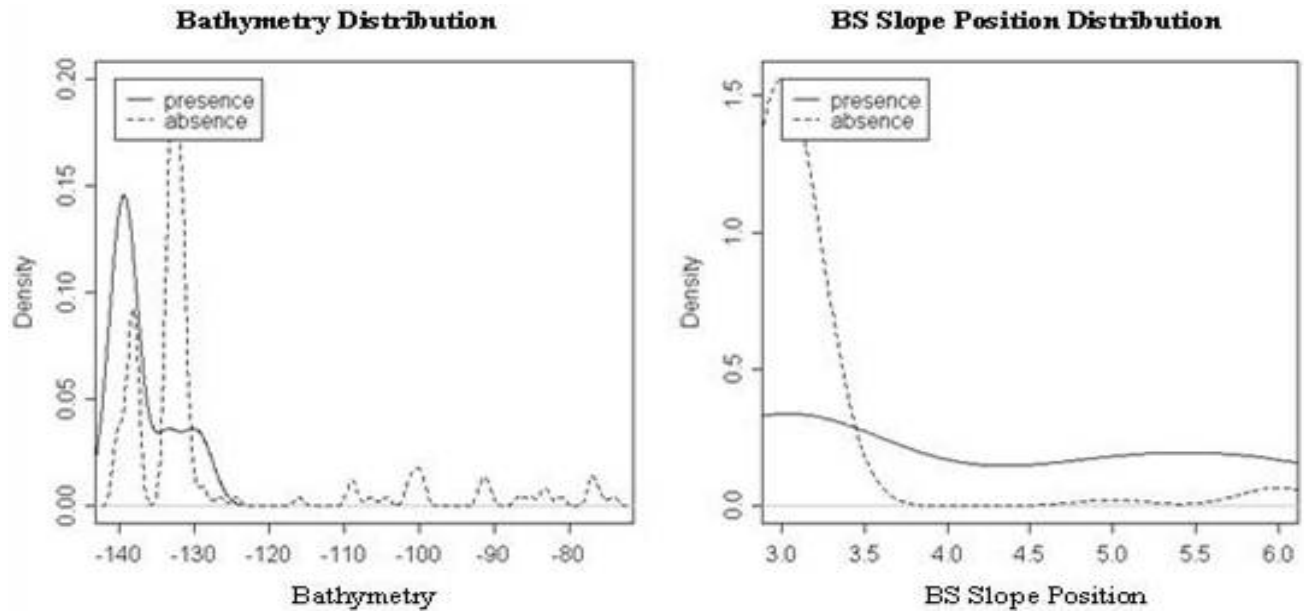
Figure 23: Histograms for each of the variables used in the GLM to predict the probability of occurrence of *S. elongatus* in Area 2. The densities of the presence and absence of *S. elongatus* are compared across the values of the predictor rasters.

The GLM formula for *S. elongatus* in Area 3 is:

$$-52.868 - 0.353(\text{bathymetry}) - 2.262(\text{BS slope position})$$

This GLM predicted that it is unlikely to find *S. elongatus* in 72% of the habitat in Area 2 and that only 20% of the area was likely to contain *S. elongatus* (Table 27, Figure 24).

Table 27: Summary of the probability results from the GLM performed on Area 2 to predict the probability of finding *S. elongatus* throughout the block.

| Probability of Occurrence	Area (m^2)	Area (km^2)	Percentage of Area 2
0.0-0.1	11813526	11.813	59%
0.11-0.2	1145160	1.145	6%
0.21-0.3	756153	0.756	4%
0.31-0.4	574722	0.570	3%
0.41-0.5	497700	0.497	2%
0.51-0.6	502659	0.502	3%
0.61-0.7	523719	0.523	3%
0.71-0.8	619362	0.619	3%
0.81-0.9	864693	0.864	4%
0.91-1.0	1908945	1.908	10%
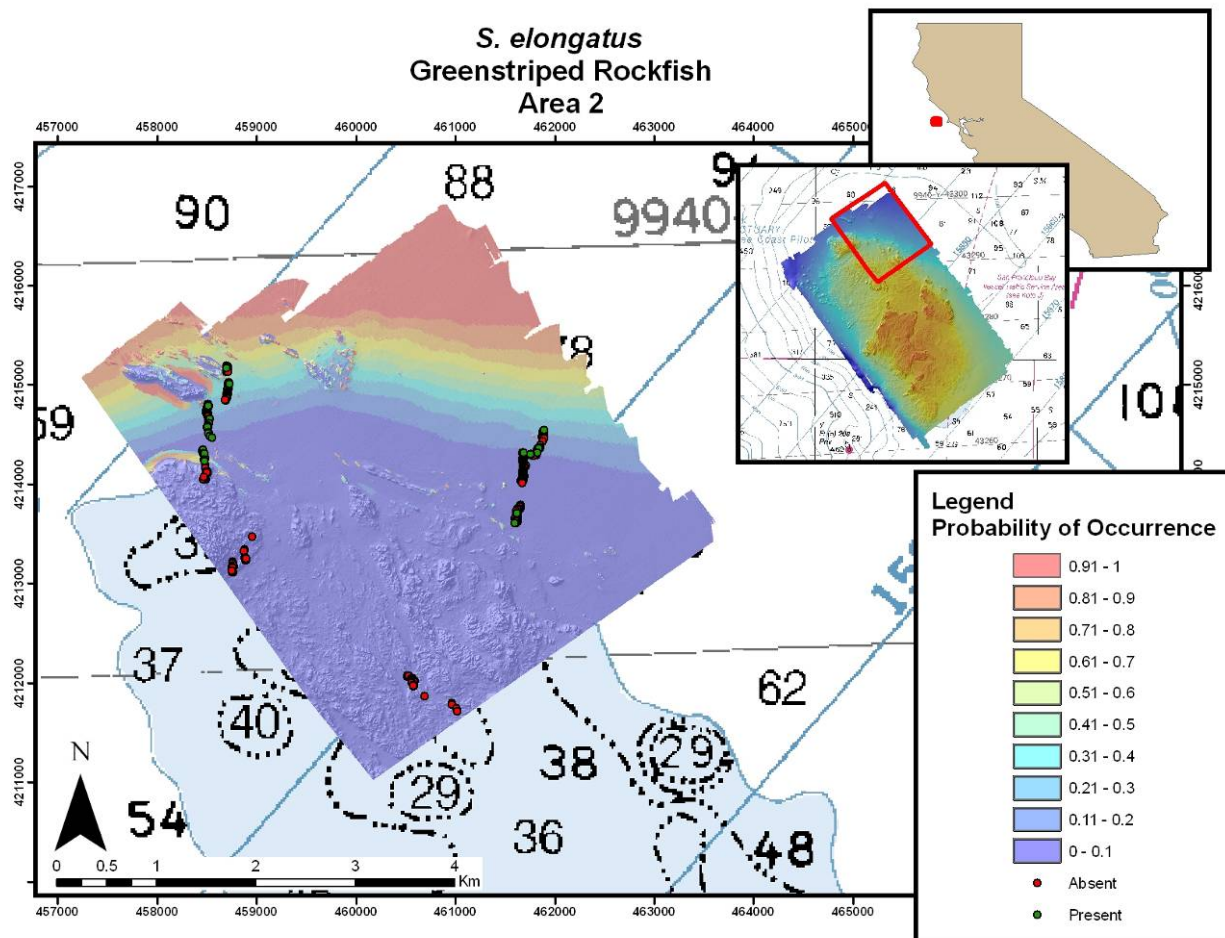
Figure 24: Area 2 GLM results for *S. elongatus* (greenstriped rockfish). Warmer colors indicate a high predicted probability of *S. elongatus* occurrence, while cooler colors indicate low probabilities. Green dots signify locations where *S. elongatus* were observed and the red dots indicate locations where no fish were present. (Image resolution: 3m; coordinate System: UTM 10N WGS 84; NOAA Chart 18640, soundings in fathoms; multibeam survey completed fall 2005.)

Since there were not enough presence points in any of the other blocks for *S. elongatus*, only GLMs were run in Areas 1 and 2. When displayed next to each other, there are obvious discrepancies between the different models. The GLM for Area 1 has very little pattern and appears mottled in its predictions while the GLM for Area 2 looks similar to a contour map (Figure 25).
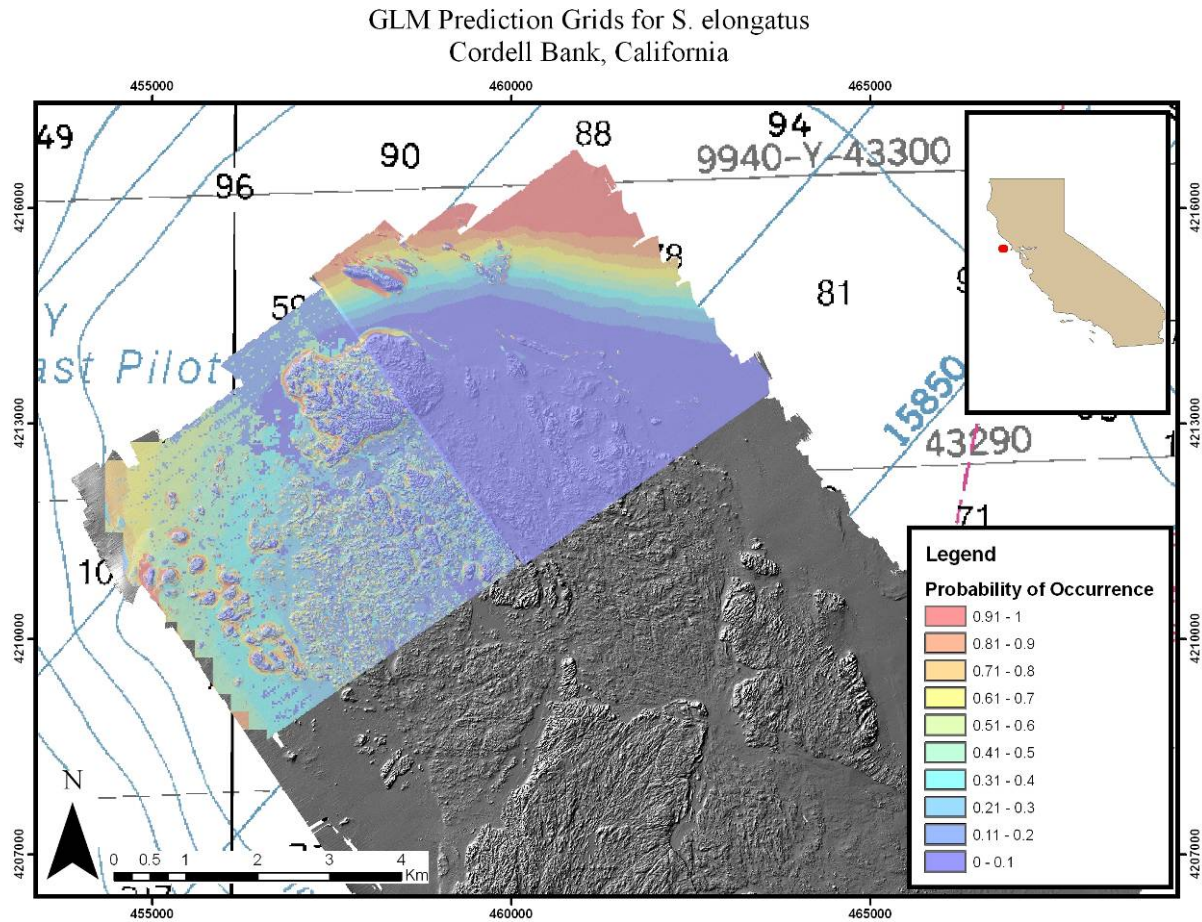
Figure 25: A combination of the probability of occurrence rasters for *Sebastes elongatus* across the entire site. Warmer colors indicate a high predicted probability of *S. elongatus* occurrence, while cooler colors indicate low probabilities. (Image resolution: 3m; coordinate System: UTM 10N WGS 84; NOAA Chart 18640, soundings in fathoms; multibeam survey completed fall 2005.)

3.6 *Submersible Video Data Analysis*

The results from the video analysis showed that the fish species chosen in this study are associated with certain depth ranges and substrate types. *S. flavidus* were found in a depth range from 50m to 140m. The majority of the yellowtail rockfish were found between 70m and 100m depth (Figure 26). The depth range for *S. rosaceus* was 50-130m, with the majority of the fish found between 50m and 70m depths (Figure 27). *Sebastes elongatus* were found slightly deeper in a depth range from 100m to 200m with some sightings around 255m (Figure 28).
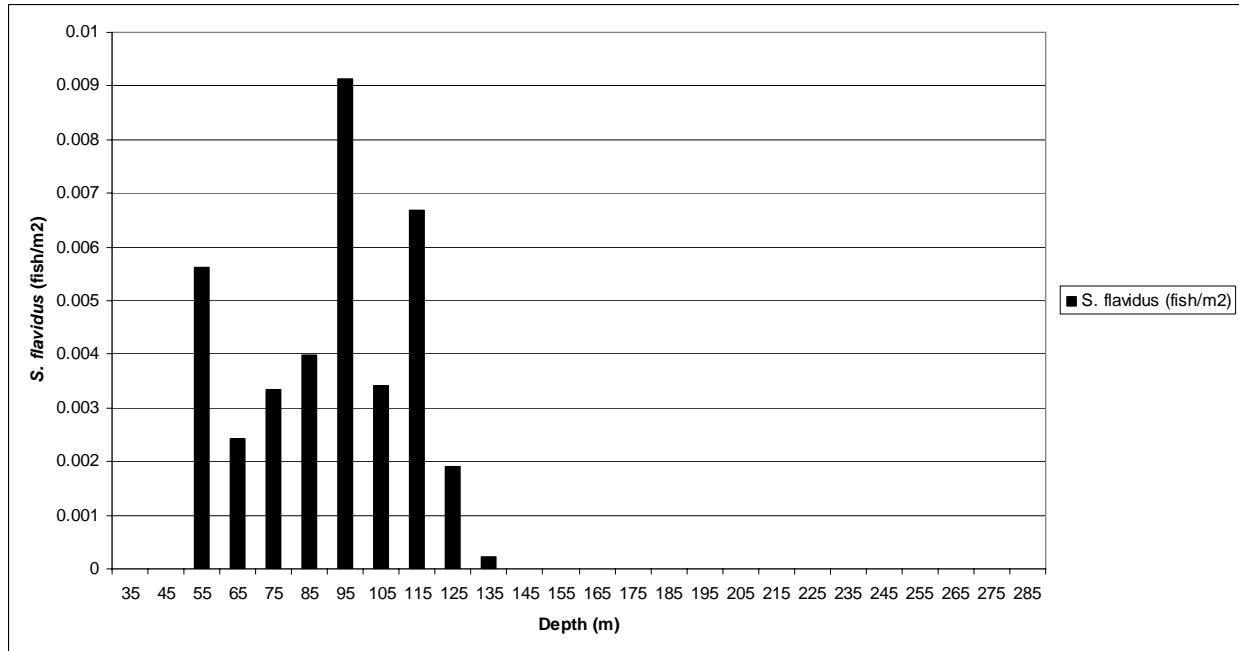
Figure 26: Density of *S. flavidus* observed at each depth from 35m to 285m on Cordell Bank, California (n = 1303).
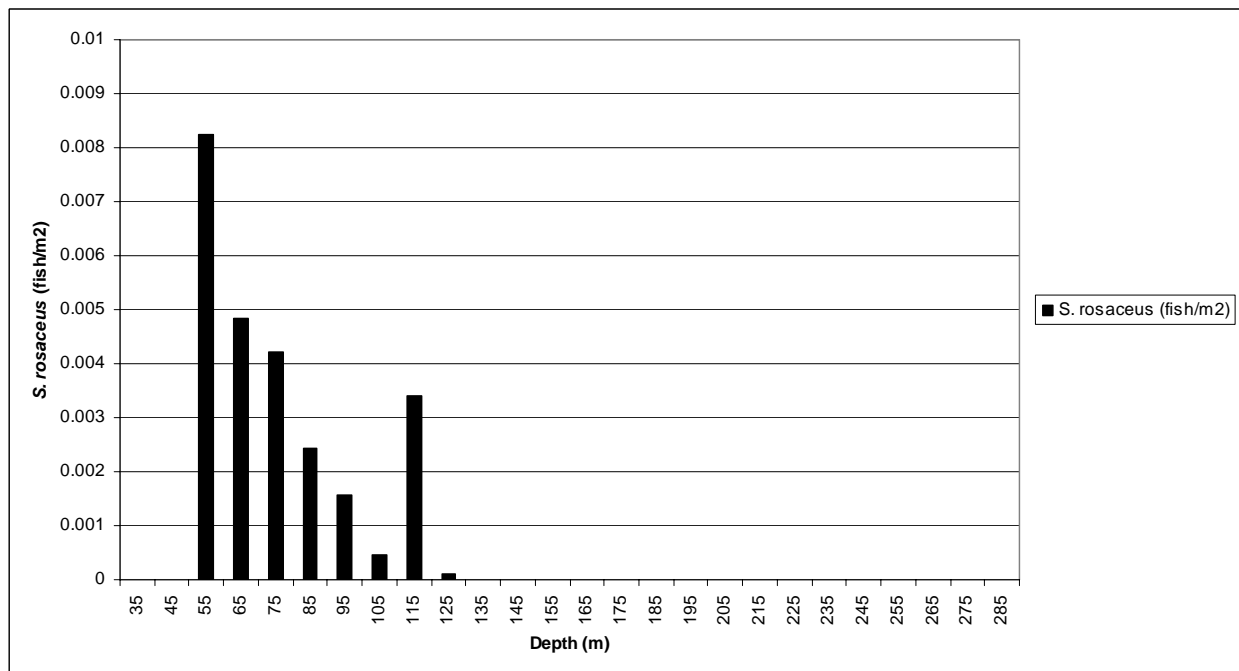
Figure 27: Density of *S. rosaceus* observed at each depth from 35m to 285m on Cordell Bank, California (n = 998)

53
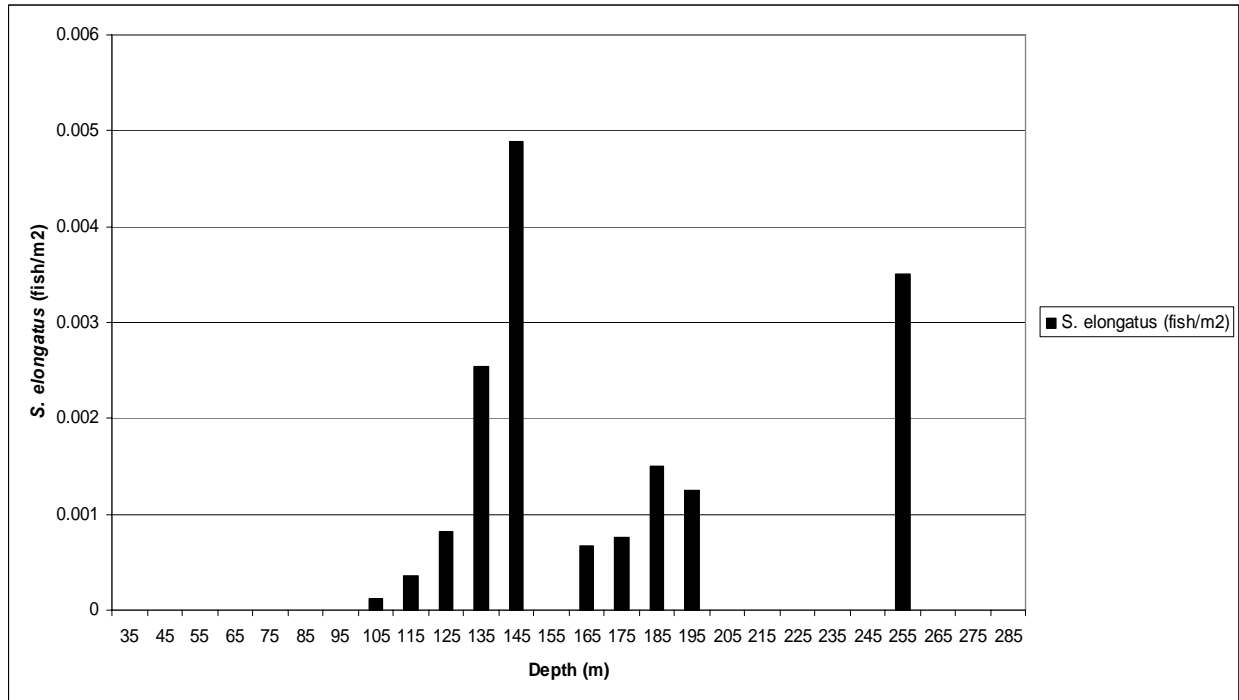
Figure 28: Density of *Sebastes elongatus* observed at each depth from 35m to 285m on Cordell Bank, California (n = 100).

The density of *S. flavidus* compared between the training transects and the evaluation transects followed similar patterns. The majority of the fish observed were found in areas of rock (r) or boulder (b), the higher relief habitats. Few fish were observed in soft sediment such as mud (m) or sand (s) (Figures 29 and 30).
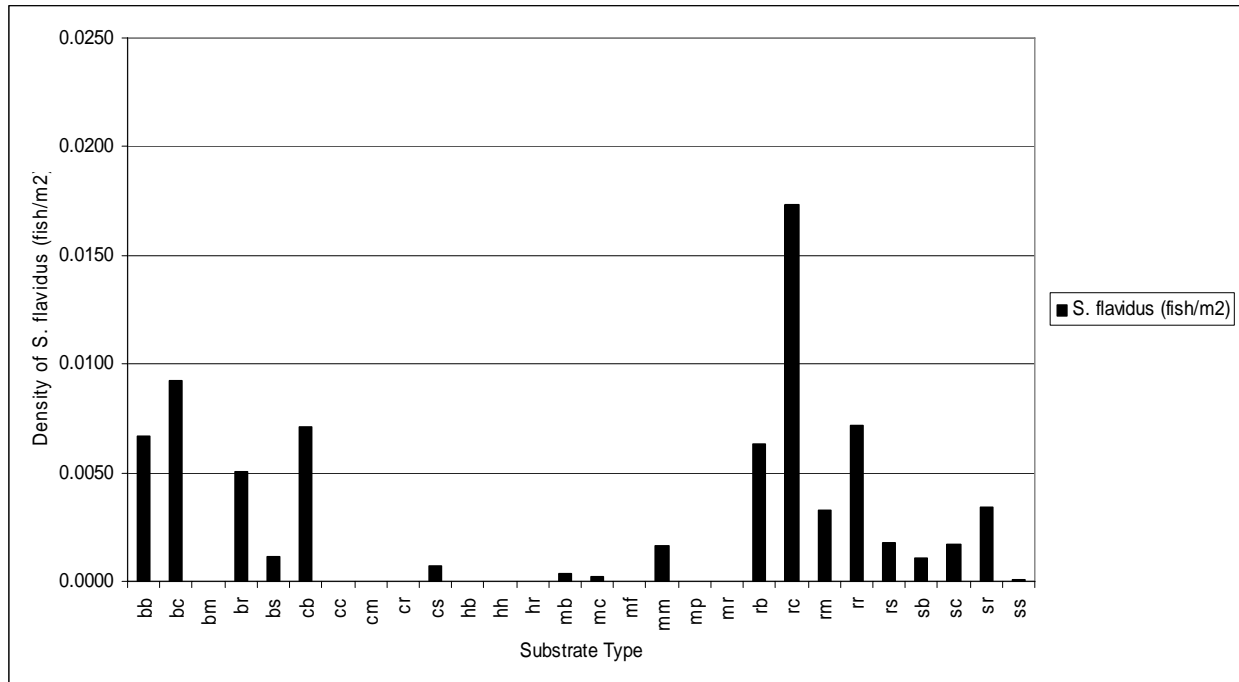
Figure 29: Density of *S. flavidus* observed over each substrate type classified from the video data within the training transects on Cordell Bank, California (Pirtle, 2005). For habitat codes, refer to Table 2 (n = 906)
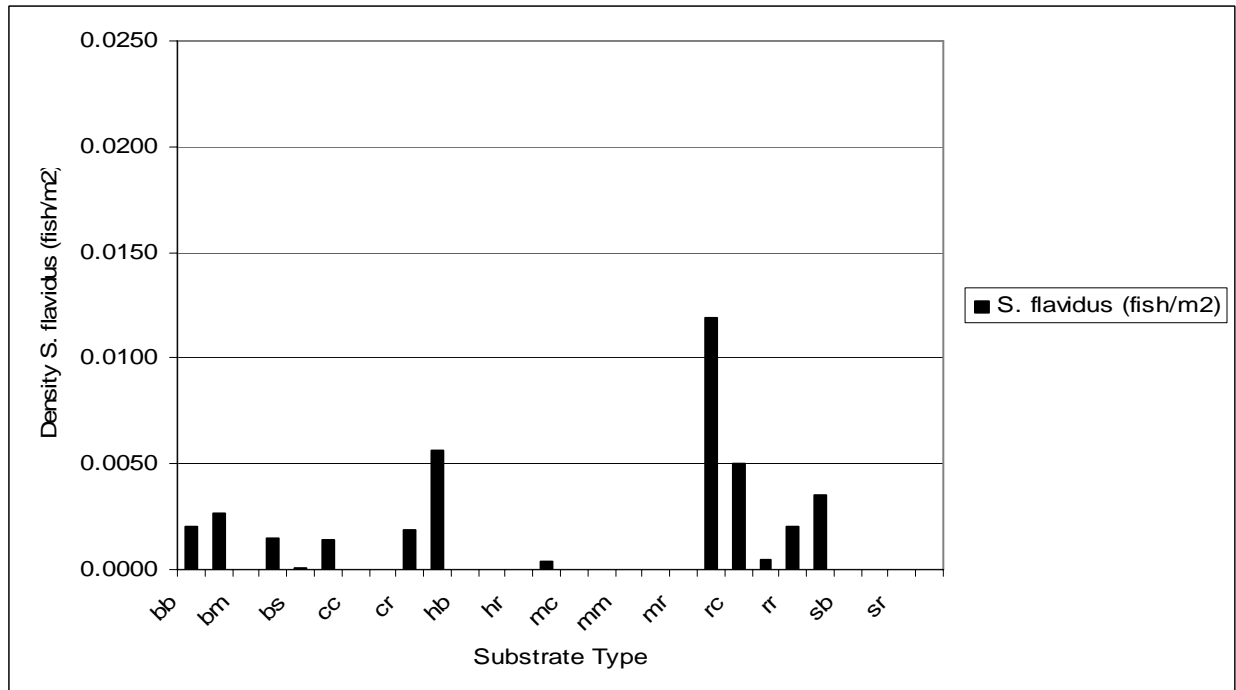
Figure 30: Density of *S. flavidus* observed over each substrate type classified from the video data within the evaluation transects on Cordell Bank, California (Pirtle, 2005). For habitat codes, refer to Table 2 (n = 397).

The density of *S. rosaceus* followed similar distribution patterns between the training and evaluation transects. Like *S. flavidus*, *S. rosaceus* are mainly found in high-relief habitats such as rock (r) and boulder (b). Some of the fish in the training transects were found in mud habitats but none of the fish were observed over the muddy substrates throughout the evaluation transects (Figures 31 and 32).
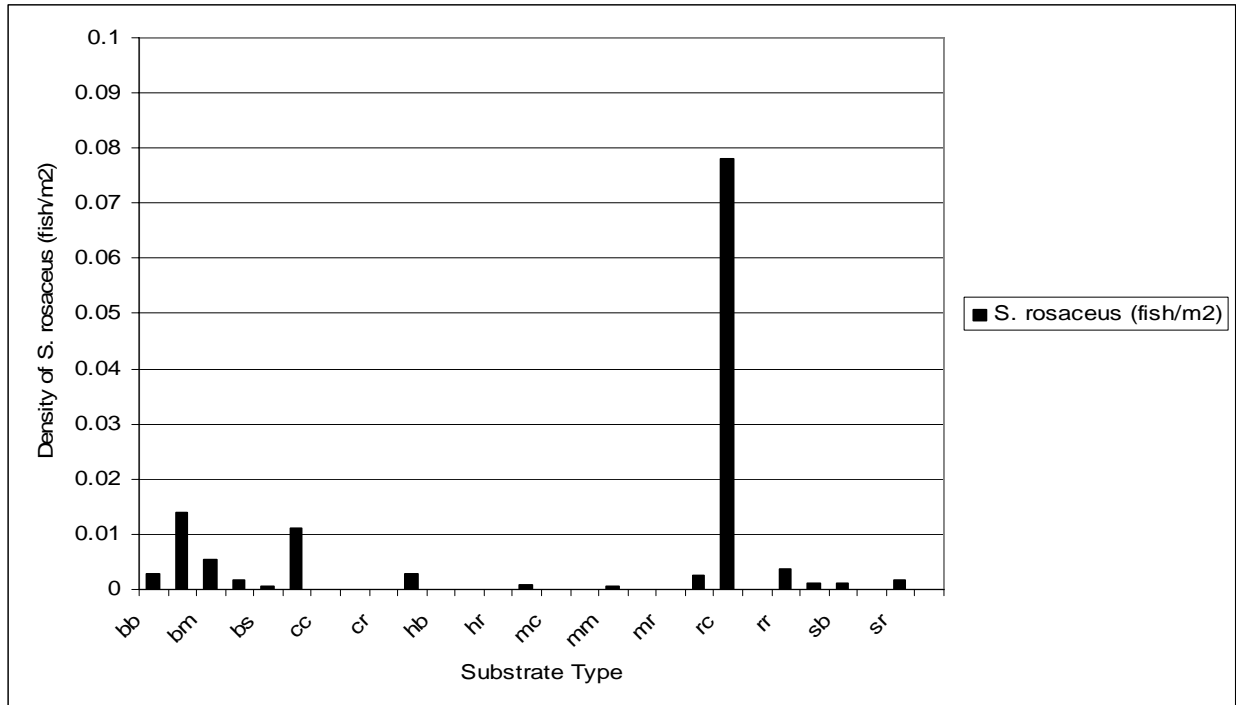
Figure 31: Density of *S. rosaceus* observed over each substrate type classified from the video data within the training transects on Cordell Bank, California (Pirtle, 2005). For habitat codes, refer to Table 2 (n = 488).
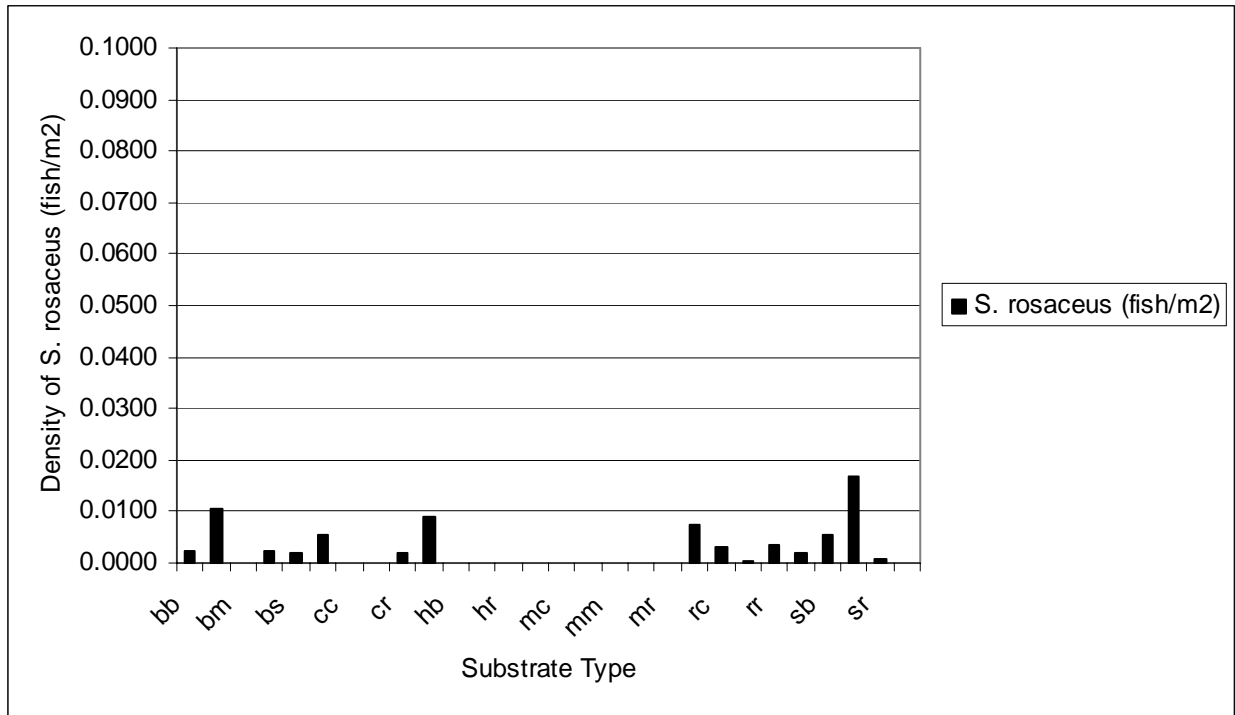
Figure 32: Density of *S. rosaceus* observed over each substrate type classified from the video data within the evaluation transects on Cordell Bank, California (Pirtle, 2005). For habitat codes, refer to Table 2 (n = 510).

The majority of *S. elongatus* were found over muddy substrates (m). Only a few fish were found over cobble (c) or boulder substrates in both the training and evaluation transects (Figures 33 and 34).
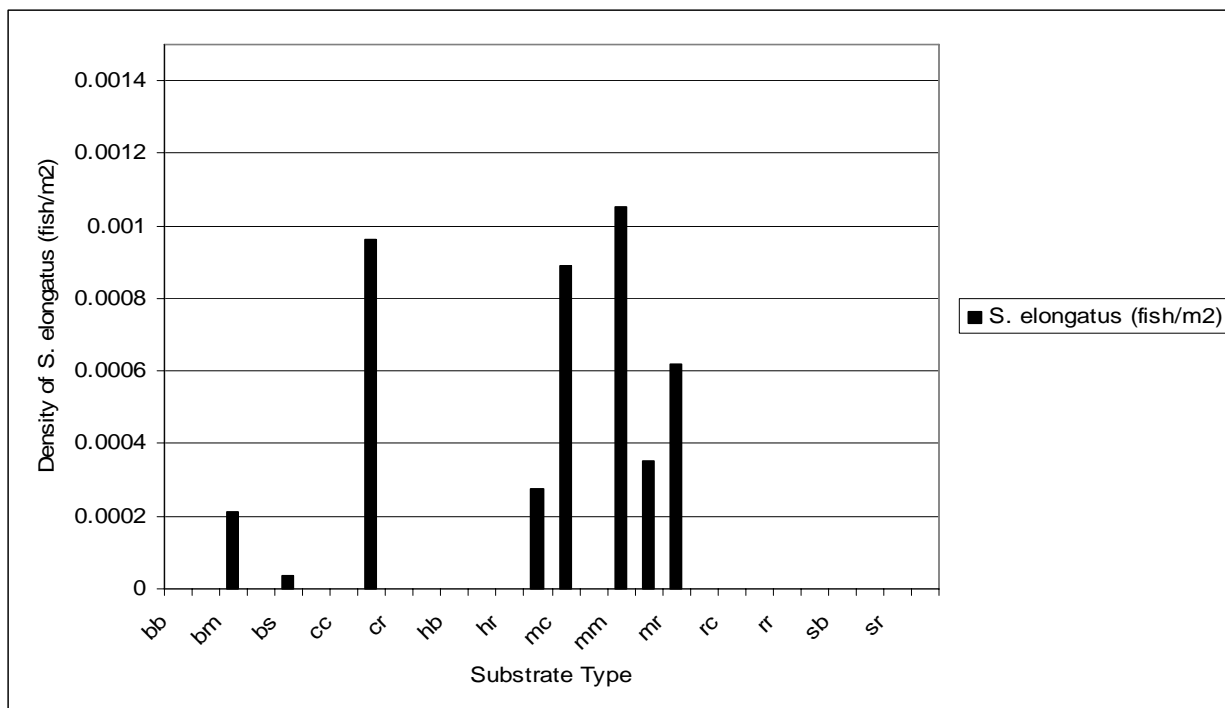
Figure 33: Density of *S. elongatus* observed over each substrate type classified from the video data within the training transects on Cordell Bank, California (Pirtle, 2005). For habitat codes, refer to Table 2 (n = 50).
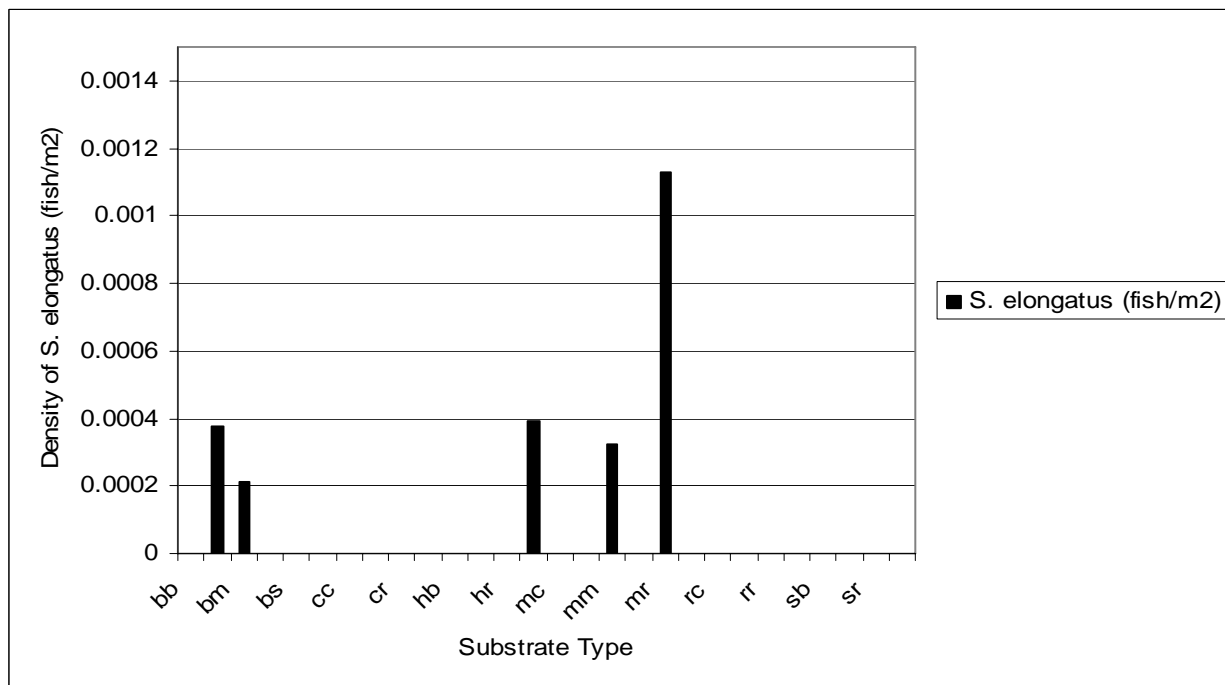
Figure 34: Density of *S. elongatus* observed over each substrate type classified from the video data within the evaluation transects on Cordell Bank, California (Pirtle, 2005). For habitat codes, refer to Table 2 (n = 50).

3.7 *Comparison of Video Analysis versus GLM Predictions*

The comparison of the predictions from the video analysis versus the predictions from the GLMs, showed that the GLMs were, on average, better at predicting the abundance of *Sebastes flavidus* and *S. rosaceus* in the eight representative transects chosen for this analysis. The percent error between the predicted and actual abundance of *S. flavidus* was 16% from the GLM predictions and 68% from the video predictions (Table 28).

Table 28: The predicted abundance of *S. flavidus* using results from the video analysis and results from the GLM analyses compared to the actual number of *S. flavidus* observed in each corresponding transect.

Transect	# of *S. flavidus* predicted using Video Analysis	# of *S. flavidus* predicted from using GLMs	Actual Number of *S. flavidus* observed in Transect
1	45	89	76
5	60	23	9
7	3	2	1
11	63	21	13
15	20	0	0
21	78	56	64
23	23	18	11
32	10	0	6
Total	**302**	**209**	**180**
Percent Error	**68%**	**16%**	**0%**

The percent error between the number of *S. rosaceus* predicted by the GLMs in each transect and the actual number observed was 13% while the percent error between the number of *S. rosaceus* predicted by the video analysis and the actual number observed was 44%. Therefore, the GLMs were more accurate at predicting the abundance of *S. rosaceus* (Table 29).

Table 29: The predicted abundance of *S. rosaceus* using results from the video analysis and results from the GLM analyses compared to the actual number of *S. rosaceus* observed in each corresponding transect.

Transect	# of *S. rosaceus* predicted using Video Analysis	# of *S. rosaceus* predicted from using GLMs	Actual Number of *S. rosaceus* observed in Transect
1	19	27	22
5	64	33	28
7	6	4	0
11	25	34	38
15	5	0	0
21	28	23	17
23	11	9	10
32	8	0	0
Total	**166**	**130**	**115**
Percent Error	**44%**	**13%**	**0%**

The percent error from the video analysis and GLM predictions were the same for *S. elongatus*. Both predicted that there would be 19 fish in the training transects when there were actually 24. Therefore, the percent error for both techniques was -21% (Table 30).

Table 30: The predicted abundance of *S. elongatus* using results from the video analysis and results from the GLM analyses compared to the actual number of *S. elongatus* observed in each corresponding transect.

Transect	# of *S. elongatus* predicted using Video Analysis	# of *S. elongatus* predicted from using GLMs	Actual Number of *S. elongatus* observed in Transect
1	0	9	0
5	0	0	0
7	1	0	0
11	0	0	0
15	12	10	13
21	0	0	0
23	0	0	0
32	6	0	11
Total	**19**	**19**	**24**
Percent Error	**-21%**	**-21%**	**0%**

3.8 *Stock Assessments*

The stock sizes calculated for each species of fish showed that *S. flavidus* has the highest predicted stock of the three species modeled on Cordell Bank. The total stock estimation for S. flavidus is 469,447 fish with the greatest number of fish in Area 1 (270,895 +/- 100,955). The stock size estimation for *S. rosaceus* across the whole site was 295,003 (+/- 65,139) with its highest estimation in Area 1 of 110,538. Compared to the other two species, the predicted stock size of *S. elongatus* was relatively small. Only 37,452 (+/- 15,906) individuals of *S. elongatus* are predicted to reside on Cordell Bank, California (Table 31, Figure 35).

Table 31: Summary table of the stock assessment predictions for *S. flavidus*, *S. rosaceus*, and *S. elongatus* on Cordell Bank, California. Errors were calculated by multiplying the percentage of inaccurate results for each species of fish in each block by the total stock.

Block	# of *S. flavidus*	# of *S. rosaceus*	# of *S. elongatus*
Area 1	270,895 ± 40,634	110,538 ± 23,213	23,231 ± 12,777
Area 2	53,894 ± 10,779	8,790 ± 791	14,221 ± 3,129
Area 3	47,139 ± 26,398	57,097 ± 21,126	n/a
Area 4	38,851 ± 10,878	48,978 ± 9,796	n/a
Area 5	50,025 ± 12,006	42,231 ± 8,024	n/a
Area 6	8,642 ± 259	27,368 ± 2,189	n/a
Total	**469,447** ± 100,955	**295,003** ± 65,139	**37,452** ± 15,906
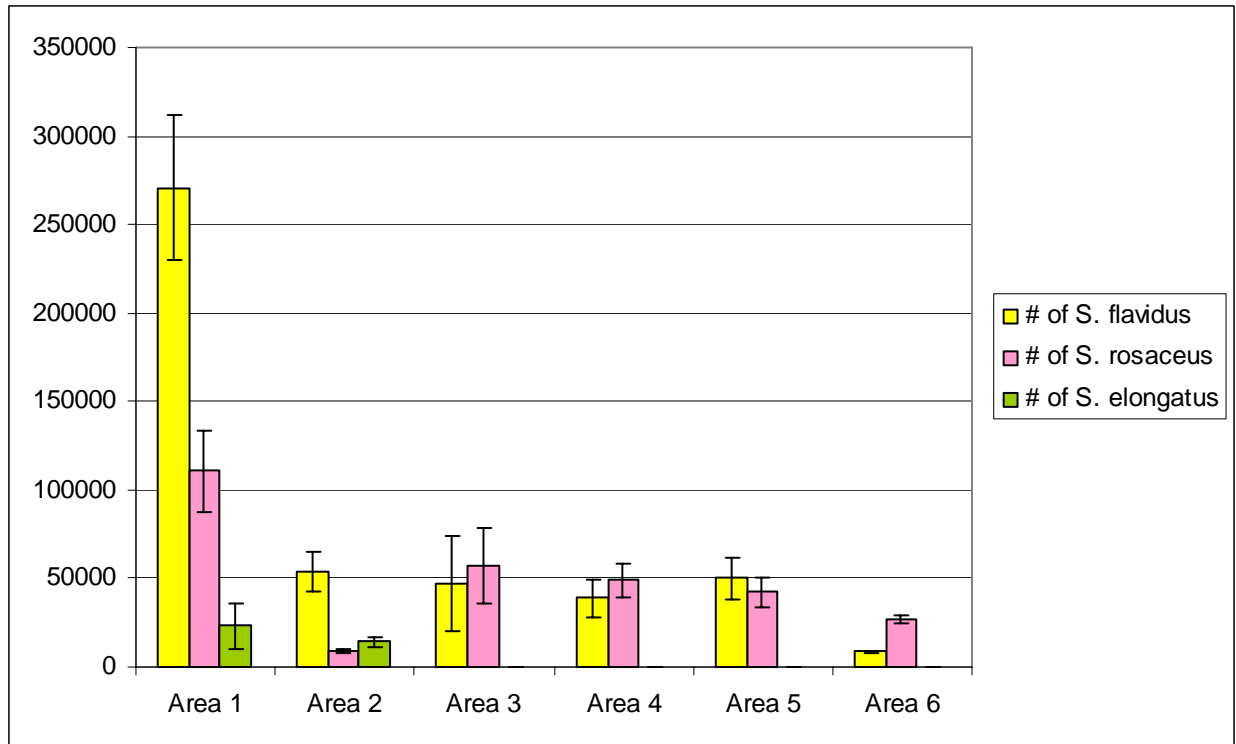
Figure 35: Total stock calculated for *S. flavidus, S. rosaceus*, and *S. elongatus* within each block of Cordell Bank, California. Error bars were calculated by multiplying the percentage of inaccurate results for each species of fish in each block by the total stock.

4 DISCUSSION

One of the most difficult aspects of fishery management is determining the current stock sizes for different species of fish. Since most stock assessments require the use of bottom trawls, accurate assessments for species that live in rocky habitats, such as rockfish, are often precluded from this technique. Since many species of rockfish are closely associated with certain habitat parameters, it is likely that stock data can be acquired through the use of habitat data, such as seafloor bathymetry and its derived products. This study attempted to determine the stock sizes of three different species of rockfish (*S. flavidus*, *S. rosaceus*, and *S. elongatus*) on Cordell Bank, California. Since, in most cases, habitat could be used to model the abundance and distribution of rockfish with some degree of accuracy, the null hypothesis was rejected.

4.1 *Multibeam Slope Analysis*

Since the categorization of the slope grid combined the majority of the values into one class ("sloping"), the classified slope grid did not come out significant in any models. Therefore,

the raw slope values were used as inputs into the GLMs rather than the categorized values. However, even the raw slope values did not prove to be very good predictors of rockfish distribution and abundance. It was only found significant in one of the 14 models run on the different species.

Although slope did not work well in any of the models from this study, it should not be disregarded as unimportant to rockfish habitat. Rockfish may not be directly related to slope, but they could be related to some derivative of slope such as the standard deviation of slope or $(slope)^2$. These options should be explored before classifying slope as an ineffective predictor of rockfish distribution and abundance.

4.2 *Multibeam Substrate Analysis*

As with slope, the binary substrate grid from the multibeam did not come out as significant in any models. However, the raw rugosity values were significant to six of the 14 models. Rugosity, the bumpiness of the terrain, appears to be a good variable to delineate between the different types of rocky substrate. From looking at the histograms displaying rugosity, *Sebastes flavidus* and *S. rosaceus* are most commonly present in areas where the rugosity values are between 1.05 and 1.2. This type of habitat is rocky but not overly complex. These species especially do not seem to prefer the higher rugosity values where there are large-scale changes in the terrain.

4.3 *Multibeam Topographic Position Index (TPI) Analysis*

The slope position grids derived from the topographic position (TPI) analysis proved to be significant predictors of rockfish abundance and distribution in the majority of the models. *S. flavidus* and *S. rosaceus* were mostly found between the grid values of 3 (middle slope) and 5 (peak/ridge). These species tend to prefer the habitats that are of greater elevation than the surrounding environment. They were rarely associated with crevices and valleys.

On the other hand, *S. elongatus* were not significantly associated with any of the values from the slope position grids. Even when these variables were found to be significant in the models predicting the probability of occurrence of *S. elongatus*, the presence of this species was evenly distributed over all values. Therefore, the GLMs for *S. elongatus* could be more effective at predicting distribution if the slope position variables were removed. Even though they were

found to be significant, the histograms show that they may be hindering the models rather than helping increase their predictive power.

4.4 *Acoustic Backscatter Analysis and Habitat Classification*

Although the methods used in this analysis have proven to be successful at classifying the habitat of sites based on the acoustic backscatter, the final product from this type of habitat classification could not be used to predict the abundance and distribution of rockfish species in this study. The final product was a vector dataset and the GLMs can only accept raster data. However, it would be interesting to develop another method to associate rockfish with this habitat classification to determine how effective it would be at predicting their abundance and distribution compared to the other methods explored in this study.

Even though the final product from this analysis was not used for the rockfish/habitat associations, one of the intermediate products, the maximum likelihood classification (MLC), was the most significant factor in one of the models for *S. elongatus*. Since the sidescan mosaic could not be utilized due to processing limitations, the MLC provided the next best alternative for differentiating between the low-relief habitats. However, if more processing power could be achieved, it would be useful to use some of the other intermediate products such as the homogeneity and entropy classifications because they would pick up on the fine scale variations in sediment type.

4.5 *Rockfish/Habitat Associations Analysis Using GLMs*

The habitat parameters used were: depth, slope, rugosity, topographic position index (TPI), aspect, and seafloor substrate. In all the GLMs created to determine the probability of occurrence for each individual species, depth was a significant factor. Depth was the only variable that was used in every model, regardless of species type, and it was often the most significant factor. As shown through the analysis of the video data, *S. flavidus*, *S. rosaceus*, and *S. elongatus* are most common in a fairly narrow depth range.

Although depth is a significant factor in all the models, it cannot accurately predict the distribution of rockfish without other habitat parameters such as rugosity. Depth does not take into account the characteristics of the seafloor and, therefore, the GLMs require at least one other habitat parameter to predict the probability of occurrence of rockfish. For example, *S. flavidus*

may prefer to remain in depths between 55m and 135m but they also prefer habitat with an adequate amount of relief. Both of these parameters have to be represented in the model in order to create an accurate distribution map of this species.

After depth, the slope position and rugosity rasters appeared as significant in the most models. Slope, aspect, and substrate were significant in a few models but not as often as the other variables. Although the maximum likelihood classification (MLC) of the acoustic backscatter mosaic was included in every model, it was only found to be significant for one of the models for *S. elongatus*. Since *S. elongatus* are commonly found to be associated with low-relief, muddy habitats, it is logical that they would find those predictors that differentiate between different substrates to be more significant.

The results from these models show that the distribution of *S. rosaceus* appears to be the easiest to model. The GLMs for *S. rosaceus* had higher levels of performance and experienced greater accuracies. Although the adjusted D^2 value for *S. flavidus* were relatively low compared to *S. rosaceus*, the models still were fairly accurate when tested with the presence and absence points from the evaluation transects. On the other hand, the GLMs for *S. elongatus* did not seem to perform very well. Although Area 2 had a high accuracy, that is most likely due to the large amount of area within the block that was assigned very low probabilities of occurrence. This caused the absence points, which there were more of, to fall into areas of low probabilities, making the model seem accurate.

The reason that the GLMs for *S. rosaceus* and *S. flavidus* were more accurate is most likely due to the fact that these species are found in rocky habitats as opposed to *S. elongatus*, which are found in muddy habitats. With the variables that were able to be included in the models (depth, rugosity, slope, aspect, TPI, and substrate) it is easier to differentiate between different rocky habitats than it is to differentiate between subtle differences in the mud habitat that could help to better model the preference of *S. elongatus*. The acoustic backscatter mosaic would probably allow for stronger predictions of *S. elongatus* probabilities of occurrence but backscatter could not be included in the GLMs because there were too many differences in values and the models would crash due to limitations in computer processing power.

Another potential cause of the poor models for *S. elongatus* could be the small amount of observations that occurred for this species. The great inequality in presence versus absence points may have caused the inaccurate results. In both areas, there were far more absence points

than there were presence points. In order to decrease the inequality of presence versus absence points, the absence points could be sub-sampled to an amount not more than twice the presence points so that not all are being used in the models. This would decrease the amount of absence points and still have representation over the areas where *S. elongatus* are not present. Also, different methods could be applied to those fish that have very few presence points such as ordination techniques. This method is based on reciprocal averaging of species and site scores and is better at dealing with data that has many absences or zeros (Guisan and Zimmerman, 2000). This type of analysis was not explored in this study.

Since the rockfish/habitat associations analysis in this study could not be run on the entire site with the use of all the fish presence/absence points, the site had to be subdivided into 6 separate parts. This introduced some inconsistency in the results. Areas that had a lot of fish points spread out over a greater number of habitat parameters seemed to make better predictions and allowed for the use of more habitat parameters in the models. For example, Area 6 only had transects that were run in one corner of the block, leaving the majority of the area under-represented. Therefore, the model predicted that the fish would have higher probabilities of being found in that corner because there were no presence points away from there. That doesn't necessarily mean that the fish do not prefer habitat away from that corner but that is the way the model interpreted the data that was available.

This study only explored the use of generalized linear models (GLMs) in order to determine the probability of occurrence of the rockfish species. However, another type of model that could potentially provide more robust predictions is the generalized additive model (GAM). GAMs are similar to GLMs in that they can deal with data that is non-linear and has non-constant variance (Guisan et al., 2002; Yee and Mitchell, 1991). However, GLMs do not always have the ability to approximate the true regression surface; especially if the true curve is asymmetrical. GLMs fit their regression surface to a curve that is symmetric and bell-shaped (Gaussian logit curve). On the other hand, GAMs are more data-driven than model-driven and, therefore, allow for a variety of different curves to be modeled (Yee and Mitchell, 1991). They are especially useful when binary data is used but can be applied to all the different types of data that GLMs are used for (Yee and Mitchell, 1991). Since the data used in this study was binary (presence/absence) it is possible that GAMs could have created more accurate models. However, GAMs could not be tested in this study because they require the use of more memory

for their processes and limitations in current equipment did not allow for this option to be explored.

Another technique that could be investigated would be including not just the rasters with raw values but also some function of those values. For example, certain species of fish may not be directly related to raw slope values but could be directly related to the slope values squared. Therefore, it would be interesting to create a model that includes functions of values as well as the raw values to see if some of the functions are more significant in the model.

In addition, new methods to determine the accuracy and fit of the models would be useful. The exploration of more rigorous statistical tests is needed to make better comparisons between models. For example, the use of a contingency table to determine the accuracy is sometimes misleading such as with *S. elongatus* in Area 2. Since the majority of the site had low probability of occurrence for this species and the number of absence points was far greater than the number of presence points, the results showed that this model was 78% accurate. Statistical tests that take into account these types of situations would allow for better evaluation of the predictions made by these models.

Finally, although the type of model used in this study was developed and has been heavily explored in the terrestrial realm, there may need to be some alterations to the methods before applying them to the marine realm. Different environmental processes occur in water than on land and these processes may need to be taken into account. Also, most of the success in previous terrestrial studies has focused on the distribution of different types of vegetation, which are sessile organisms. Fish are not sessile organisms, therefore, fine-tuning the models to take into account not just the habitat attributes that are directly beneath the presence/absence points but also those that are in very close proximity could help to improve the effectiveness of the GLMs by recognizing that fish do not remain in one spot.

4.6 *Submersible Video Data Analysis*

The analysis performed on the rockfish from the submersible video data outlined some of the major patterns in the depth distribution and habitat associations of *S. flavidus, S. rosaceus*, and *S. elongatus*. However, unlike the GLMs, the results from the video analyses cannot be extrapolated to encompass the whole site. Only habitat where video data was collected can be defined if other imagery is not available such as maps produced from acoustic surveys to

associate the video classification with. In addition, the process used to distinguish between the different types of habitats is fairly subjective. For example, one person may define a certain type of habitat as cobble while someone else labels it as boulder. Finally, a great amount of time has to be allotted to watching the videos and classifying the habitat compared to the amount of time it takes to run one of the analyses on the multibeam or acoustic backscatter data within GIS to differentiate between the different habitats.

4.7 *Comparison of Video Analysis versus GLM Predictions*

The comparison between the video analyses versus the GLM predictions showed that, for this study, the GLMs were better at predicting the "stock" sizes of *S. flavidus* and *S. rosaceus* than the habitat characterization from the video analysis. For *S. flavidus*, the percent error between the predicted and observed stock size was only 16% compared to a percent error of 68% from the video analysis predictions. The percent error from the video analysis predictions of the stock sizes of *S. rosaceus* was better than for *S. flavidus* (44%) but it still was greater than the percent error from the GLM predictions (13%). Therefore, the GLMs were more accurate at making predictions for these two species of fish that are often associated with rocky habitats.

On the other hand, there was no difference in the percent error between the video analysis and the GLM analysis for S. elongatus. They both had a percent error of -21%. However, the GLM analysis could only utilize 3 of the 8 test transects used in this comparison because no models existed under the other 5 transects. Therefore, this may not be an accurate comparison because the percent error from the video analysis was for all 8 transects.

Based on these results, the use of GLMs seems to be more effective at predicting the stock of *S. rosaceus* and *S. flavidus*. In addition to being more accurate, the GLM predictions can also be extrapolated over a greater area than the video data. Video data alone cannot be used to classify habitat across the entire site because only the areas where there is video data can be classified using the techniques utilized in this study. Also, the video analysis only took into account substrate type and it did not consider other factors such as depth. Since depth was such an important predictor of species distribution in the GLMs, it can be assumed that the video analysis is not very effective because of the lack of depth data. Therefore, for two of the species, the GLMs proved to be more effective and efficient. In further studies, however, it would be

useful to come up with a statistical test to determine whether or not these differences in accuracy are significant between the two methods.

For *S. elongatus*, better models would need to be developed to more accurately predict their distribution and abundance using GLMs. At this point, it would make sense if the video data were better at predicting their abundance. Being a species that is associated with soft substrates, better predictors would need to be used to determine their preferences within the GLMs. Currently, the video analysis is better at distinguishing between these subtle differences in the low-relief habitat.

4.8 *Stock Assessments*

Although stock assessments were calculated for each species within each block, the predicted stock counts are limited by the accuracy of the GLMs. In areas where the GLMs were not very effective, the stock assessments may be greater or less than the actual stock size, depending on whether the model tends to over-predict or under-predict the abundance of fish. As a preliminary study, the stocks of fish were calculated to show that as a possible use of these models; however, more robust models would be desired before using them to predict stock sizes. For example, the stock assessment of *S. elongatus* is probably far from the actual stock size because those models were not very accurate in their predictions.

Even with the limitations of the stock assessments produced in this study, the stock sizes do follow the general pattern expected with these species of fish. *S. flavidus* had the greatest predicted stock size of 469,447 fish (+/- 100,955), followed by *S. rosaceus* with 295,003 fish (+/- 65,139), and then *S. elongatus* with 37,452 (+/- 15,906). The video observations supported this pattern as well. Of these three fish, *S. flavidus* was observed the most, then *S. rosaceus*, and finally *S. elongatus* with the least amount of sightings. In addition, a catch and release study done on the bank in 1994 ranked *S. flavidus* and *S. rosaceus* as the first and second most common species on the bank, respectively (Eldridge, 1994).

5 CONCLUSION

In conclusion, remote sensing data and landscape ecology analysis and modeling tools can be used to accurately predict the distributions of rock fish species. If these models prove to

be consistent and accurate after further testing, this study could provide an efficient technique for managing fishery resources all over the world.

In addition, with additional development of these methods, it may be possible to create a process that allows for stock assessments in areas where very limited biological data exists. For example, if certain species of fish are found to be consistently and strongly associated with certain habitat characteristics derived from the multibeam and backscatter data, it could be possible to simply map an area and, with some knowledge of the species of fish that are found there, predict the abundance of those species. Or they could be used to determine the distribution of their preferred habitats and potential stock size in areas already depleted by over fishing but candidates for MPA status. These methods could prove to be an efficient and effective technique for deciding on the placement of marine protected areas and marine reserves.

Computer technology, remote sensing, GIS, and statistical software are constantly advancing in the direction of creating new tools for the study of biodiversity. It is important that these tools are utilized to develop efficient techniques for studying biodiversity, especially in the marine environment where access is limited. The types of models that were used in this study have been utilized in the terrestrial realm for years by decision-makers and land-use planners to protect habitat for land-based species (Austin, 1998). The oceans, however, have by enlarged not benefited from this type of objective and statistically rigorous landscape analysis. With the amount of high-resolution bathymetric data that has been collected and that is being collected in the marine environment, it is very possible to implement the type of habitat modeling explored in this study in a number of different areas. Therefore, it is important to continue fine-tuning methods that can be used to accurately predict the abundance and distribution of species in the marine realm.

LITERATURE CITED

Austin, M.P., A.O. Nicholls, M.D. Doherty, J.A. Meyers. 1994. Determining Species Response Functions to an Environmental Gradient by means of a #-Function. Journal of Vegetation Science 5(2): 215-228.

Austin, M. 1998. An Ecological Perspective on biodiversity investigations: examples from Australian Eucalypt forests. Annals of the Missouri Botanical Garden 85(1): 2-17.

Best, B.D., S. Loarie, S. Qian, P. Halpin, D. Urban. 2005. ArcRStats – multivariate habitat Modeling with ArcGIS and R Statistical Software. Available at: <http://www.nicholas.duke.edu/geospatial/software.>

CDFG Website. 2006. Marine life protection act initiative. < http://www.dfg.ca.gov/mrd/mlpa/overview.html>. Accessed February 16, 2006.

Clark, C.W. 1996. Marine reserves and the precautionary management of fisheries. Ecological Applications 6(2): 369-370.

Cochrane, G.R., K.D. Lafferty. 2002. Use of acoustic classification of sidescan sonar data for mapping benthic habitat in the Northern Channel Islands, California. Continental Shelf Research 22: 683-690.

Cordell Bank National Marine Sanctuary Cruise Report. 2002. NOAA Ocean Service.

Cordell Bank National Marine Sanctuary Website. 2007. Natural Environment: Biological Resources. <http://cordellbank.noaa.gov/environment/bio_res.html>. Accessed February 16, 2006.

Eldridge, E.B. 1994. Hook-and-Line fishing study at Cordell Bank, California, 1986-1989. NOAA Technical Memorandum NMFS. NOAA-TM-NMFS-SWFSC-197.

Greene, H.G., M.M. Yoklavich, R.M. Starr, V.M. O'Connell, W.W. Wakefield, D.E. Sullivan, J.E. McRea Jr., G.M. Cailliet.. 1999. A classification scheme for deep seafloor habitats. Oceanologica Acta. 22(6): 663-678

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

Guisan, A. and N.E. Zimmerman. 2000. Predictive Habitat distribution models in ecology. Ecological Modelling. 135: 147-186.

Iampietro, P.J., R.G. Kvitek, E. Morris. 2005. Recent Advances in Automated Genus-Specific Marine Habitat Mapping Enabled by High-Resolution Multibeam Bathymetry. Marine Technology Society Series 39(3): 83-93.

Intelmann, S.S. and G.R. Cochrane. 2006. Benthic habitat mapping in the Olympic Coast National Marine Sanctuary: Classification of side scan sonar data from survey HMPR-108-2002-01: Version I. Marine Sanctuaries Conservation Series ONMS-06-01.

Jaberg, C. and A. Guisan. 2001. Modelling the distribution of bats in relation to landscape structure In a temperate environment. The Journal of Applied Ecology. 38(6): 1169-1181.

Jenness, J. 2005. Random point generator (randpts.avx) extension for ArcView 3.x, v. 1.3. Jenness Enterprises. Available at: http://www.jennessent.com/arcview/random_points.htm.

Lauck, T., C.W. Clark, M. Mangel, G.R. Munro. 1998. Implementing the precautionary principle in fisheries management through marine reserves. Ecological Applications 8(1): S72-S78.

Leslie, H., M. Ruckelshaus, I.R. Ball, S. Andelman, H.P. Possingham. 2003. Using siting algorithms in the design of marine reserve networks. Ecological Applications 13(1) Supplement: S185-S198.

Love, M.S., M. Yoklavich, L. Thorsteinson. 2002. The Rockfishes of the Northeast Pacific. Berkeley: University of California Press. 404 p.

Nasby-Lucas, N.M., S.G. Merle, B.W. Embley, B.N. Tissot, M.A. Hixon, D.J. Wright. 2002. Integration of submersible transect data and high-resolution multibeam sonar imagery for a habitat-based groundfish assessment of Haceta Bank, Oregon. Fisheries Bulletin 100: 739-751.

Pirtle, J.L. 2005. Habitat-based assessment of structure-forming megafaunal invertebrates and fishes on Cordell Bank, California. Masters Thesis.

Starr, R.M., D.S. Fox, B.N. Tissot, G.E. Johnson, W.H. Barss. 1995. Comparison of submersible-Survey and hydroacoustic-survey estimates of fish density on a rocky bank. Fisheries Bulletin 94: 113-123.

Wakefield, W.W., C.E. Whitmire, J.E.R. Clemons, and B.N. Tissot. 2005. Fish Habitat Studies: Combining high-resolution geological and biological data. American Fisheries Society Symposium 41: 119-138.

Weber, M.L. 2002. From abundance to scarcity: A history of U.S. marine fisheries policy. Washington, DC: Island Press. 245 p.

Weiss, A.D. 2001. Topographic Positions and Landforms Analysis (Conference Poster). San Diego, CA International User Conference.

Wilson, E.O. 2002. The future of Life. New York: Vintage Books. 256 p.

Wright, D.J., E.R. Lundblad, E.M. Larkin, R.W. Rinehart. 2005. Oregon State University Davey Jones Locker Seafloor Mapping/Marine GIS Lab. (project background at http://dusk.geo.orst.edu/esri04/p1433_ron.html)

Yee, T.W., N.D. Mitchell. 1991. Generalized additive models in plant ecology. Journal of Vegetation Science. 2(5): 587-602.

Yoklavich, M.M., G.H. Greene, G.M. Cailliet, D.E. Sullivan, R.N. Lea, M.S. Love. 1999. Habitat associations of deep-water rockfishes in a submarine canyon: an example of natural refuge. Fisheries Bulletin 98: 625-641.

APPENDIX

Sebastes flavidus (Yellowtail Rockfish)

AREA 1

Contingency table of the data used to determine the probability accuracy from the GLM performed in Area 1 to predict the probability of finding *S. flavidus* throughout the block (n = 227).

Point Type	# of Points where model indicates		Total
	Low Probability	High Probability	
Present	9	121	130
Absent	72	24	96
Total	82	145	227

Summary of the predictor variables used in the GLM performed on Area 1 to predict the probability of finding *S. flavidus* throughout the block.

Coefficients:	Estimate	Std. Error	z value	Pr(>|z|)	
(Intercept)	19.65906	3.962127	4.962	6.99E-07	***
Rugosity	-8.56848	3.44214	-2.489	0.0128	*
Bathymetry	0.066406	0.007645	8.686	<2.00E-16	***
BS Slope Position	-0.33039	0.136631	-2.418	0.0156	*

Summary of the probability results from the GLM performed on Area 1 to predict the probability of finding *S. flavidus* throughout the block.

Probability of Occurrence	Area (m^2)	Area (km^2)	Percentage of Area 1
0.00-0.1	606969	0.61	3%
0.11-0.2	215982	0.22	1%
0.21-0.3	148302	0.15	1%
0.31-0.4	267498	0.27	1%
0.41-0.5	412056	0.41	2%
0.51-0.6	607095	0.61	3%
0.61-0.7	1839285	1.84	9%
0.71-0.8	3348522	3.35	17%
0.81-0.9	3771180	3.77	19%
0.91-1.00	9045666	9.05	45%
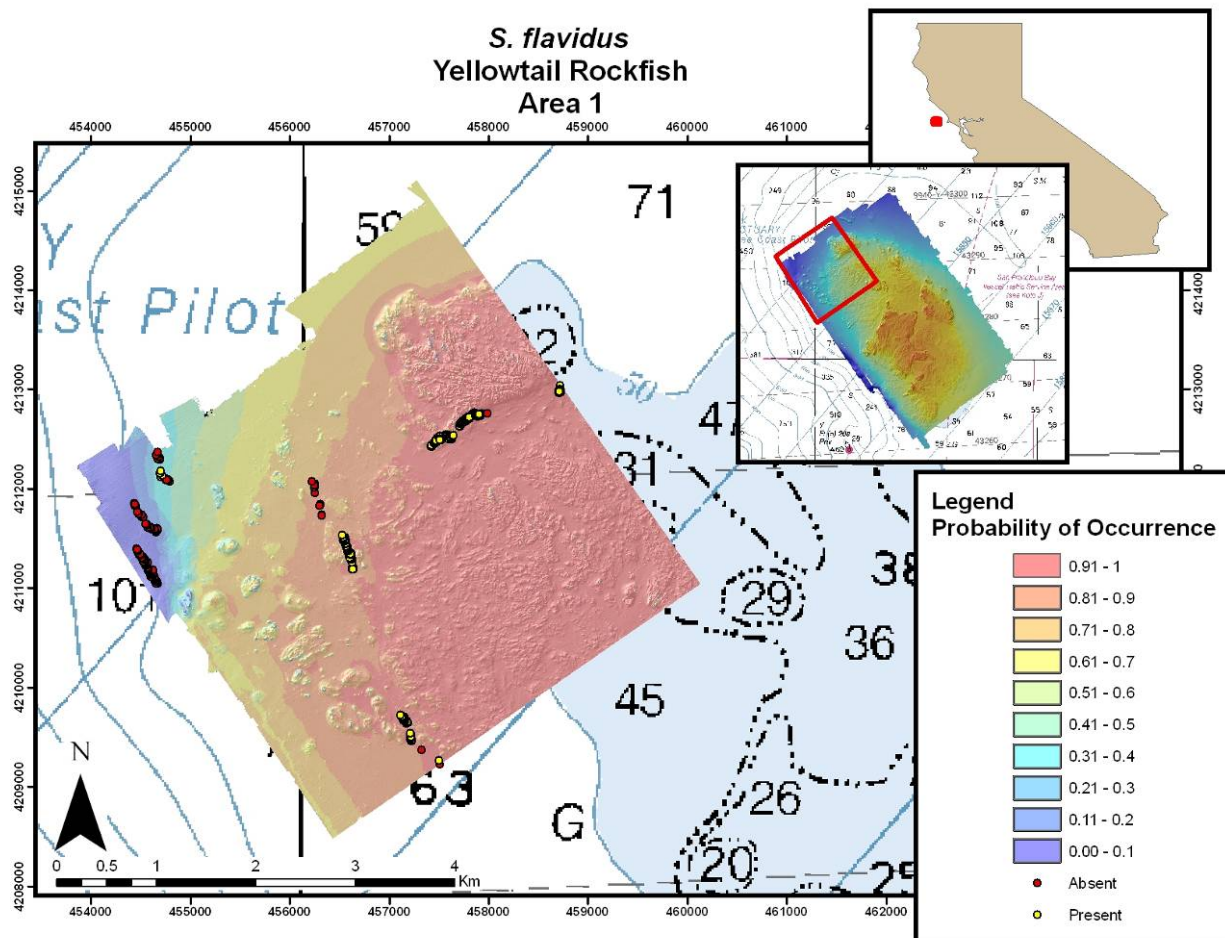
Area 1 GLM results for *Sebastes flavidus* (yellowtail rockfish). Warmer colors indicate a high predicted probability of *S. flavidus* occurrence, while cooler colors indicate low probabilities. Yellow dots signify locations where *S. flavidus* were observed and the red dots indicate locations where no fish were present. (Image resolution: 3m; coordinate System: UTM 10N WGS 84; NOAA Chart 18640, soundings in fathoms; multibeam survey completed fall 2005.)

AREA 2

Contingency table of the data used to determine the probability accuracy from the GLM
performed in Area 2 to predict the probability of finding *Sebastes flavidus* throughout the block
(n = 184)

| Point Type | # of Points where model indicates | | Total |
	Low Probability	High Probability	
Present	6	48	54
Absent	99	31	130
Total	105	79	184

Summary of the predictor variables used in the GLM performed on Area 2 to predict the
probability of finding *S. flavidus* throughout the block.

Coefficients:	Estimate	Std. Error	z value	Pr(>|z|)	
(Intercept)	24.872	5.763	4.316	1.59E-05	***
Bathymetry	0.084	0.012	6.896	5.36E-12	***
BS Slope Position	0.003	0.001	2.115	0.0344	*
Rugosity	-15.345	5.223	-2.938	0.003	**

Summary of the probability results from the GLM performed on Area 2 to predict the probability
of finding *S. flavidus* throughout the block.

Probability of Occurrence	Area (m^2)	Area (km^2)	Percentage of Area 2
0.0-0.1	4824504	4.82	25%
0.11-0.2	4724046	4.72	25%
0.21-0.3	1727802	1.73	9%
0.31-0.4	985869	0.99	5%
0.41-0.5	772425	0.77	4%
0.51-0.6	661320	0.66	3%
0.61-0.7	676143	0.68	4%
0.71-0.8	960102	0.96	5%
0.81-0.9	1773828	1.77	9%
0.91-1.0	2048004	2.05	11%
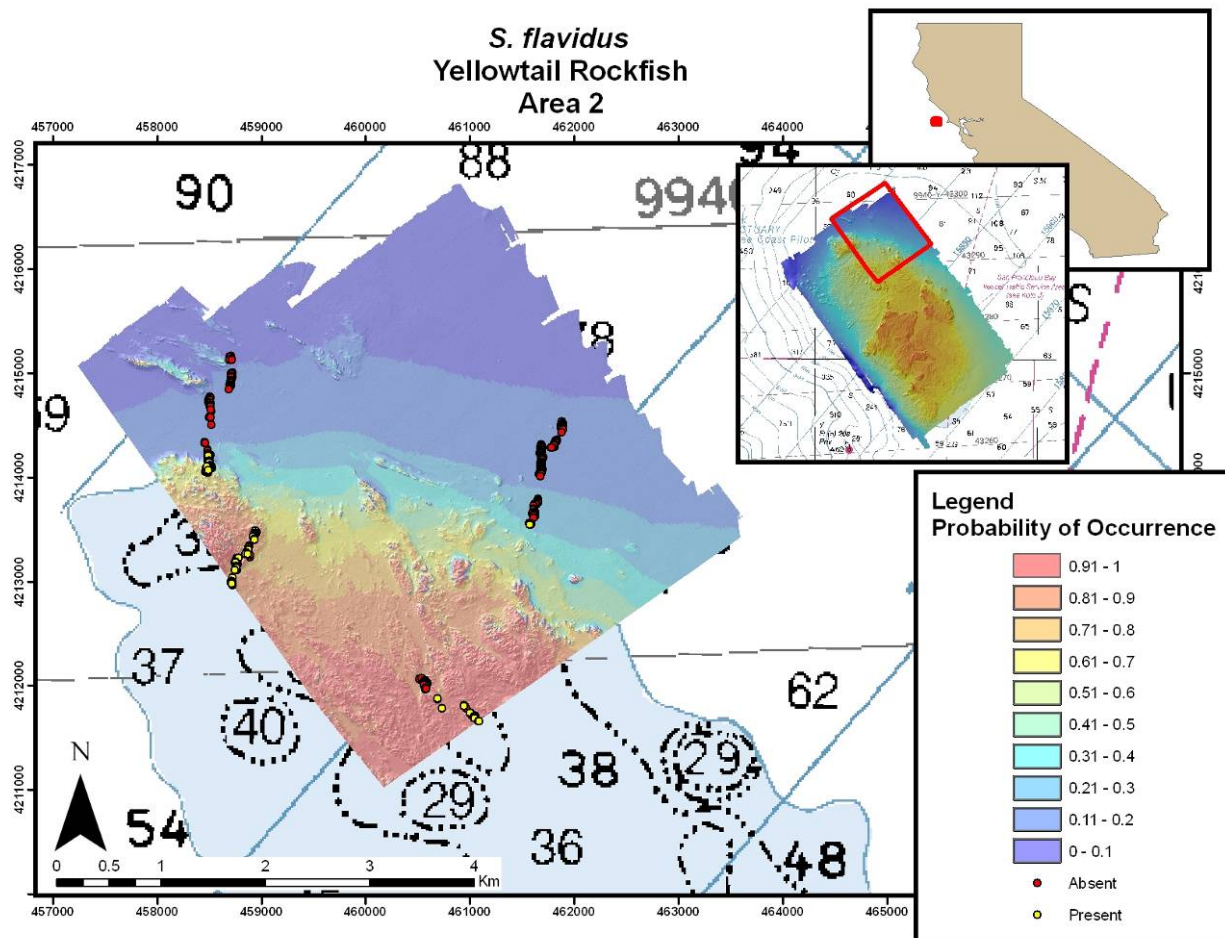
Area 2 GLM results for *Sebastes flavidus* (yellowtail rockfish). Warmer colors indicate a high predicted probability of *S. flavidus* occurrence, while cooler colors indicate low probabilities. Yellow dots signify locations where *S. flavidus* were observed and the red dots indicate locations where no fish were present. (Image resolution: 3m; coordinate System: UTM 10N WGS 84; NOAA Chart 18640, soundings in fathoms; multibeam survey completed fall 2005.)

AREA 4

Contingency table of the data used to determine the probability accuracy from the GLM
performed in Area 4 to predict the probability of finding *Sebastes flavidus* throughout the block
(n = 113).

| Point Type | # of Points where model indicates | | Total |
	Low Probability	High Probability	
Present	6	36	42
Absent	45	26	71
Total	51	62	113

Summary of the probability results from the GLM performed on Area 4 to predict the probability
of finding *S. flavidus* throughout the block.

Coefficients:	Estimate	Std. Error	z value	Pr(>|z|)	
(Intercept)	-314.453	122.931	-2.558	0.011	*
Bathymetry	0.055	0.009	5.850	4.92E-9	***
Rugosity	319.760	123.220	2.595	0.009	**
FS Slope Position	-0.923	0.335	-2.780	0.005	**
BS Slope Position	0.619	0.205	3.014	0.003	**

Summary of the probability results from the GLM performed on Area 4 to predict the probability
of finding *S. flavidus* throughout the block.

Probability of Occurrence	Area (m^2)	Area (km^2)	Percentage of Area 4
0.0-0.1	691218	0.691	3%
0.11-0.2	3771918	3.772	19%
0.21-0.3	2123586	2.124	11%
0.31-0.4	710235	0.710	4%
0.41-0.5	592497	0.592	3%
0.51-0.6	654228	0.654	3%
0.61-0.7	1739232	1.739	9%
0.71-0.8	2209797	2.210	11%
0.81-0.9	2042307	2.042	10%
0.91-1.0	6109047	6.109	31%
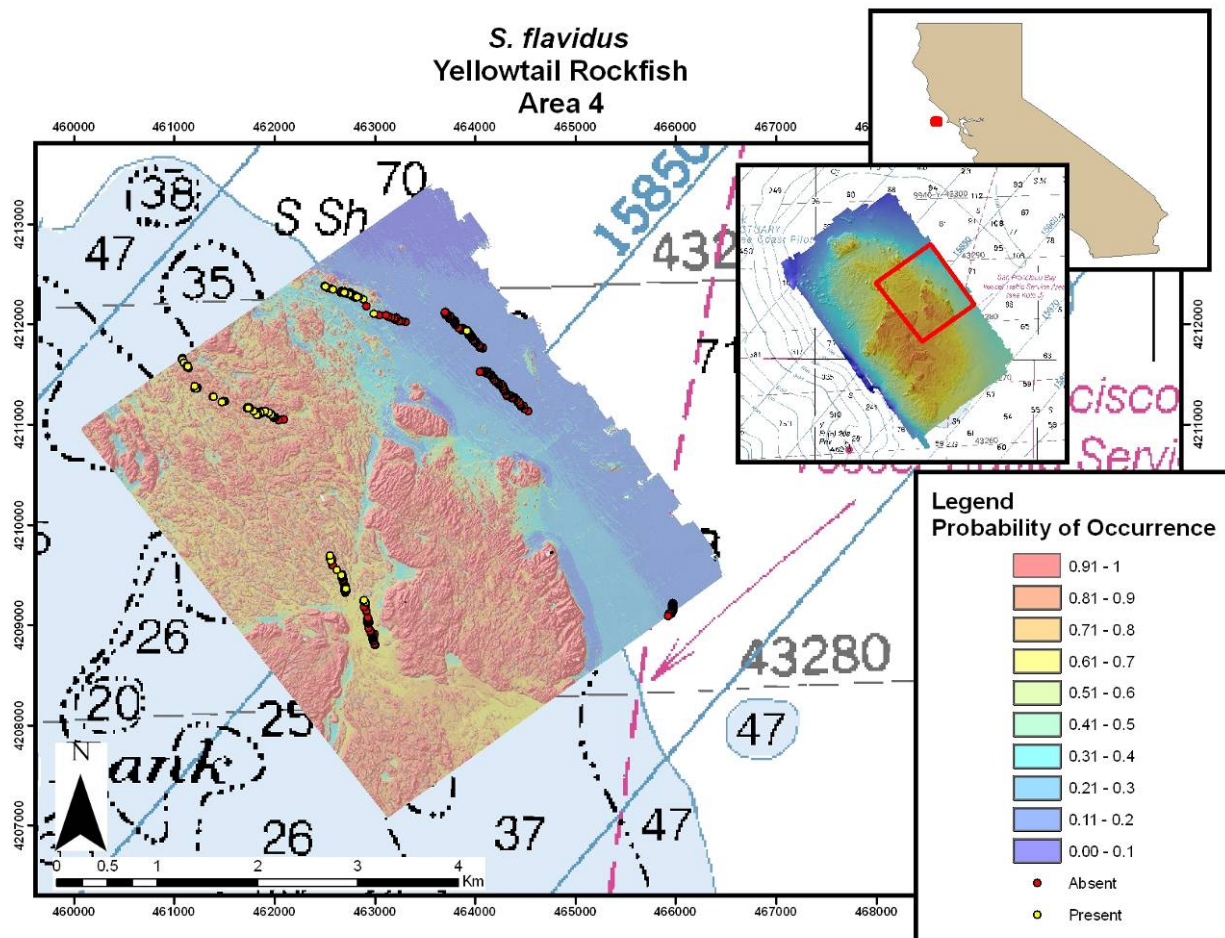
Area 4 GLM results for *Sebastes flavidus* (yellowtail rockfish). Warmer colors indicate a high predicted probability of *S. flavidus* occurrence, while cooler colors indicate low probabilities. Yellow dots signify locations where *S. flavidus* were observed and the red dots indicate locations where no fish were present. (Image resolution: 3m; coordinate System: UTM 10N WGS 84; NOAA Chart 18640, soundings in fathoms; multibeam survey completed fall 2005.)

AREA 5

Contingency table of the data used to determine the probability accuracy from the GLM performed in Area 5 to predict the probability of finding *Sebastes flavidus* throughout the block (n = 209).

| Point Type | # of Points where model indicates | | Total |
	Low Probability	High Probability	
Present	26	37	63
Absent	121	25	146
Total	147	62	209

Summary of the probability results from the GLM performed on Area 5 to predict the probability of finding *S. flavidus* throughout the block.

Coefficients:	Estimate	Std. Error	z value	Pr(>|z|)	
(Intercept)	6.689	2.571	2.602	0.009	**
Rugosity	-5.318	2.518	-2.112	0.035	*
Bathymetry	0.054	0.010	0.010	2.88E-8	***
BS Slope Position	1.073	0.213	0.213	4.92E-7	***

Summary of the probability results from the GLM performed on Area 5 to predict the probability of finding *S. flavidus* throughout the block.

Probability of Occurrence	Area (m^2)	Area (km^2)	Percentage of Area 5
0.0-0.1	1451754	1.452	7%
0.11-0.2	1365318	1.365	6%
0.21-0.3	1512972	1.513	7%
0.31-0.4	1517193	1.517	7%
0.41-0.5	1887831	1.888	9%
0.51-0.6	2397195	2.397	11%
0.61-0.7	2390832	2.391	11%
0.71-0.8	3133413	3.133	14%
0.81-0.9	2349270	2.349	11%
0.91-1.0	3632427	3.632	17%
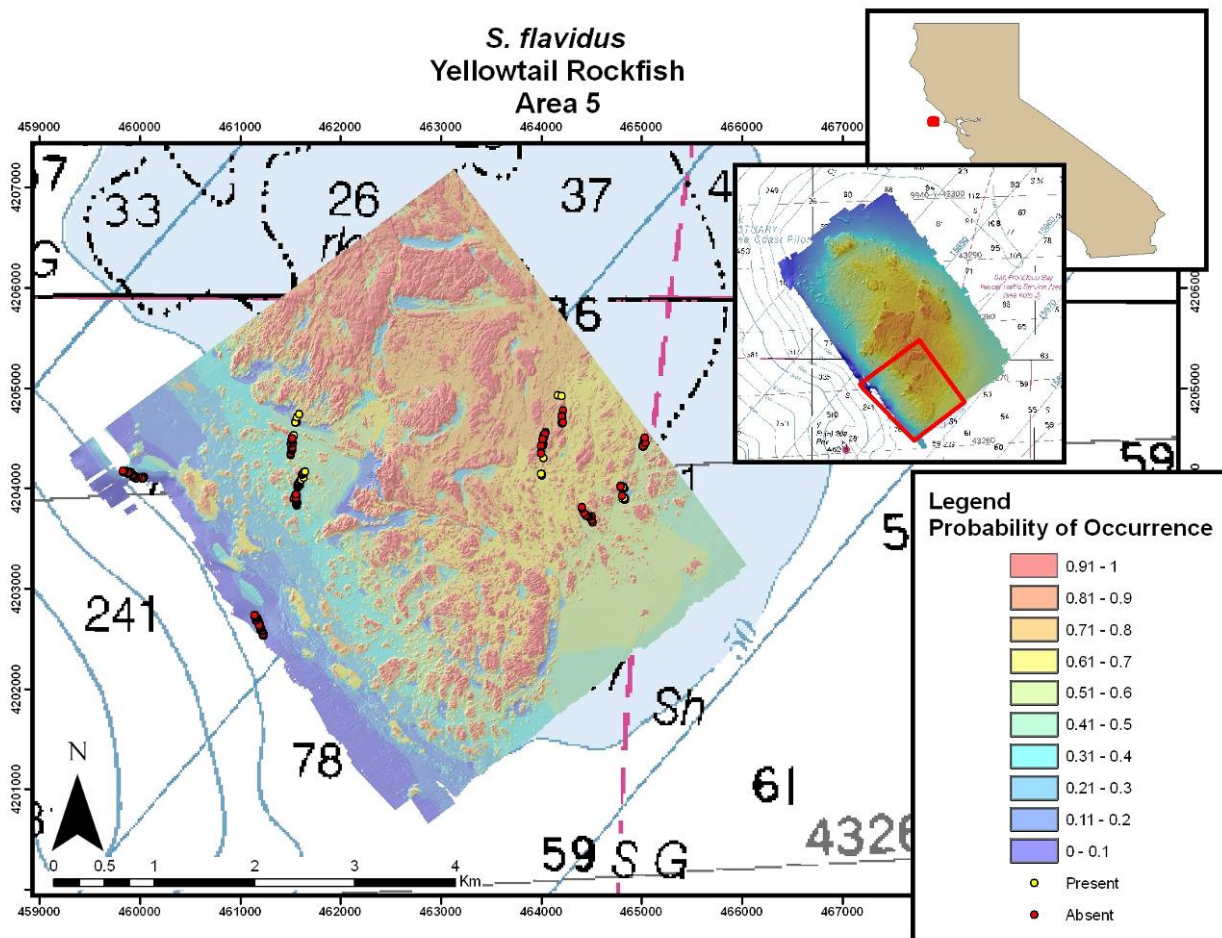
Area 5 GLM results for *Sebastes flavidus* (yellowtail rockfish). Warmer colors indicate a high predicted probability of *S. flavidus* occurrence, while cooler colors indicate low probabilities. Yellow dots signify locations where *S. flavidus* were observed and the red dots indicate locations where no fish were present. (Image resolution: 3m; coordinate System: UTM 10N WGS 84; NOAA Chart 18640, soundings in fathoms; multibeam survey completed fall 2005.)

Sebastes rosaceus (Rosy Rockfish)

AREA 1

Contingency table of the data used to determine the probability accuracy from the GLM performed in Area 1 to predict the probability of finding *S. rosaceus* throughout the block (n = 197).

Point Type	# of Points where model indicates		Total
	Low Probability	High Probability	
Present	10	66	76
Absent	90	31	121
Total	100	97	197

Summary of the predictor variables used in the GLM performed on Area 1 to predict the probability of finding *S. rosaceus* throughout the block.

Coefficients:	Estimate	Std. Error	z value	Pr(>|z|)	
(Intercept)	62.820	20.751	3.027	0.002	**
Rugosity	-47.092	18.424	-2.556	0.011	*
Bathymetry	0.096	0.036	2.659	0.008	**
BS Slope Position	-1.307	0.350	-3.724	0.000	***
FS Slope Position	0.735	0.254	2.888	0.004	**

Summary of the probability results from the GLM performed on Area 1 to predict the probability of finding *S. rosaceus* throughout the block.

Probability of Occurrence	Area (m^2)	Area (km^2)	Percentage of Area 1
0.0-0.1	1458009	1.46	7%
0.11-0.2	374184	0.37	2%
0.21-0.3	479907	0.48	2%
0.31-0.4	448587	0.45	2%
0.41-0.5	568368	0.57	3%
0.51-0.6	870084	0.87	4%
0.61-0.7	1897929	1.90	9%
0.71-0.8	2310093	2.31	11%
0.81-0.9	2711250	2.71	13%
0.91-1.0	9144153	9.14	45%
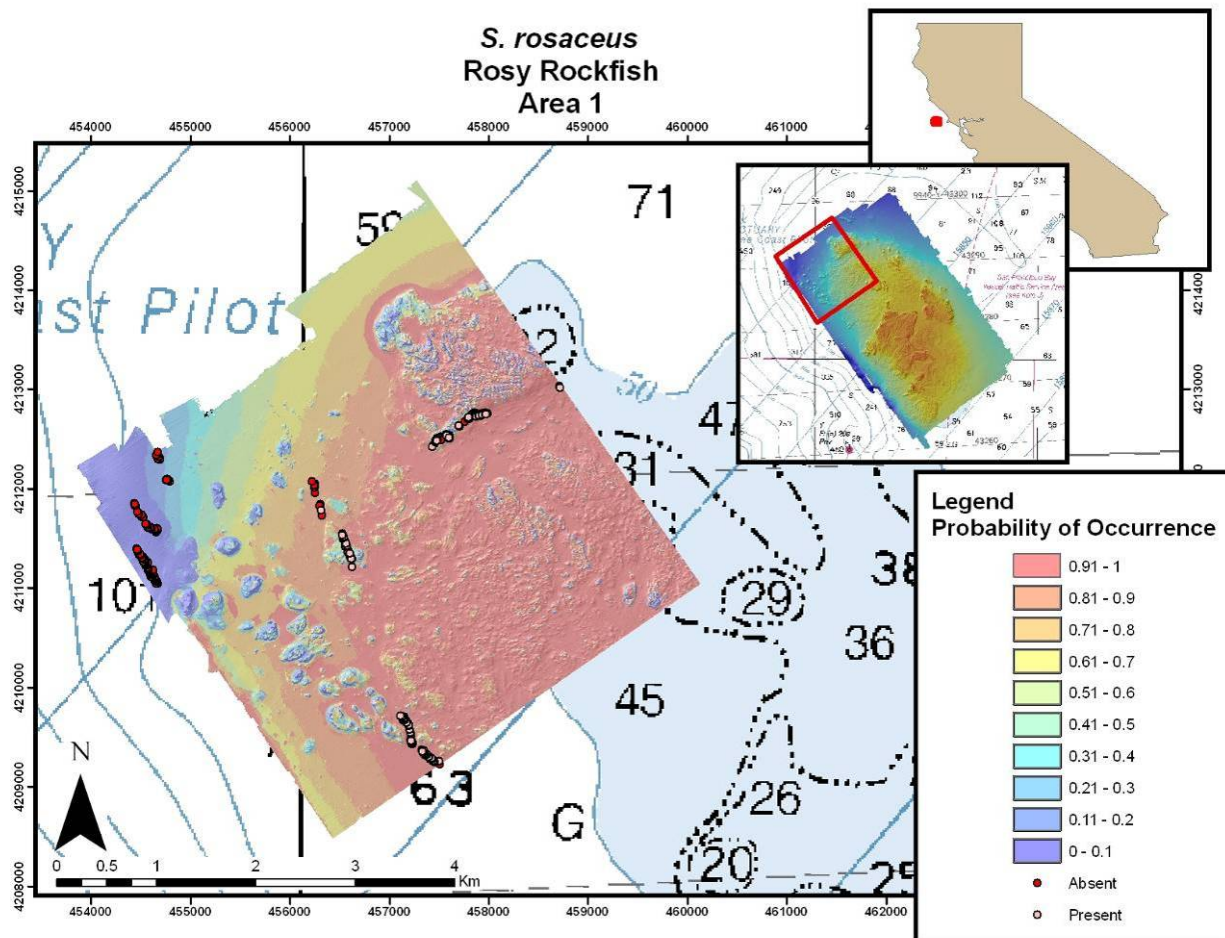
Area 1 GLM results for *Sebastes rosaceus* (rosy rockfish). Warmer colors indicate a high predicted probability of *S. rosaceus* occurrence, while cooler colors indicate low probabilities. Pink dots signify locations where *S. rosaceus* were observed and the red dots indicate locations where no fish were present. (Image resolution: 3m; coordinate System: UTM 10N WGS 84; NOAA Chart 18640, soundings in fathoms; multibeam survey completed fall 2005.)

AREA 4

Contingency table of the data used to determine the probability accuracy from the GLM performed in Area 4 to predict the probability of finding *Sebastes rosaceus* throughout the block (n = 181).

| Point Type | # of Points where model indicates | | Total |
	Low Probability	High Probability	
Present	12	22	34
Absent	123	24	147
Total	135	46	181

Summary of the probability results from the GLM performed on Area 4 to predict the probability of finding *S. rosaceus* throughout the block.

Coefficients:	Estimate	Std. Error	z value	Pr(>|z|)	
(Intercept)	10.067	3.092	3.255	0.001	**
Bathymetry	0.148	0.041	3.647	0.000	***
Slope	0.462	0.187	2.470	0.014	*

Summary of the probability results from the GLM performed on Area 4 to predict the probability of finding *S. rosaceus* throughout the block.

Probability of Occurrence	Area (m^2)	Area (km^2)	Percentage of Area 4
0.0-0.1	6998454	6.998	35%
0.11-0.2	440559	0.441	2%
0.21-0.3	312030	0.312	2%
0.31-0.4	365553	0.366	2%
0.41-0.5	566046	0.566	3%
0.51-0.6	855441	0.855	4%
0.61-0.7	1259226	1.259	6%
0.71-0.8	1556199	1.556	8%
0.81-0.9	1950588	1.951	10%
0.91-1.0	6370848	6.371	32%
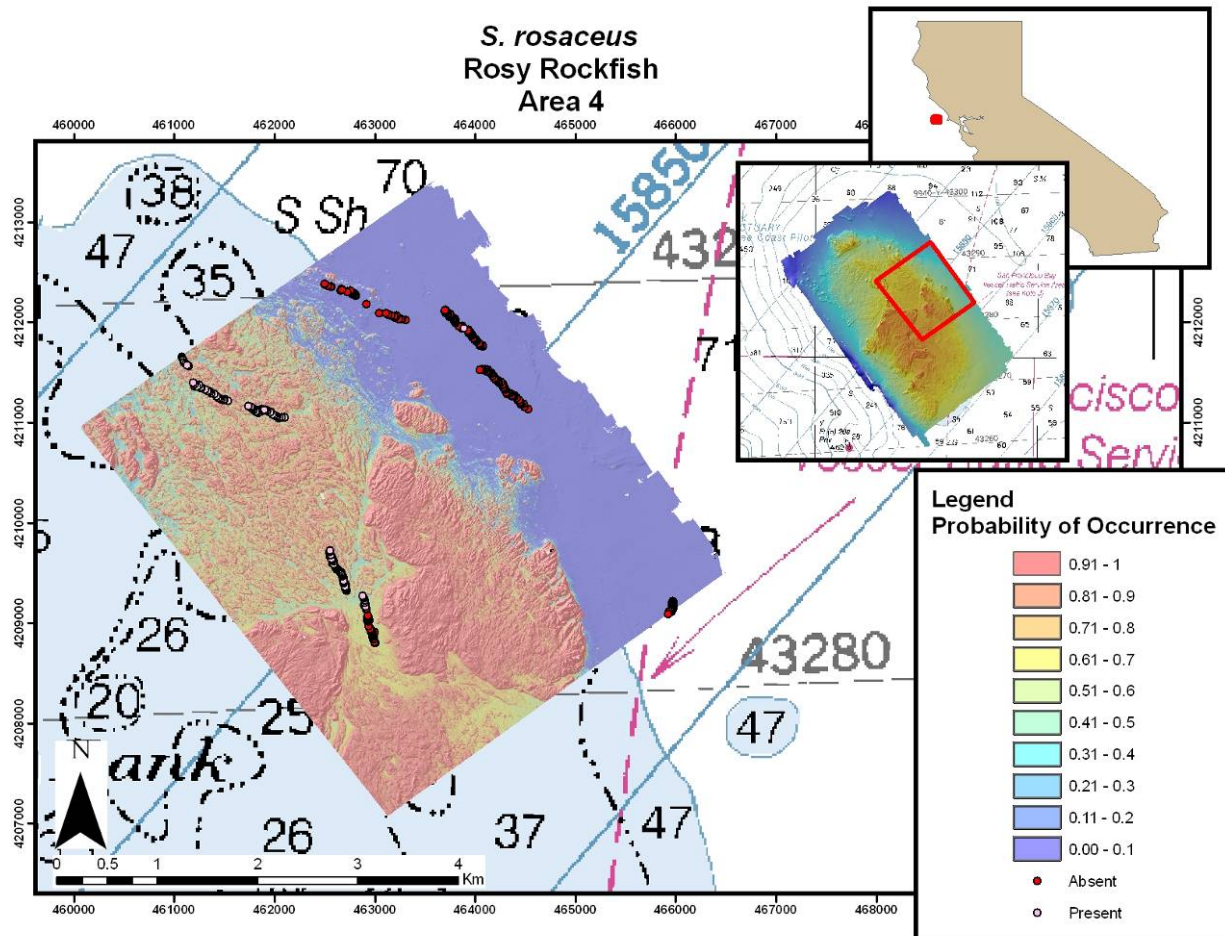
S. rosaceus
Rosy Rockfish
Area 4

Area 4 GLM results for *Sebastes rosaceus* (rosy rockfish). Warmer colors indicate a high predicted probability of S. rosaceus occurrence, while cooler colors indicate low probabilities. Pink dots signify locations where *S. rosaceus* were observed and the red dots indicate locations where no fish were present. (Image resolution: 3m; coordinate System: UTM 10N WGS 84; NOAA Chart 18640, soundings in fathoms; multibeam survey completed fall 2005.)

AREA 5

Contingency table of the data used to determine the probability accuracy from the GLM
performed in Area 5 to predict the probability of finding *Sebastes rosaceus* throughout the block
(n = 220).

| Point Type | # of Points where model indicates | | Total |
	Low Probability	High Probability	
Present	20	54	74
Absent	124	22	146
Total	144	76	220

Summary of the probability results from the GLM performed on Area 5 to predict the probability
of finding *S. rosaceus* throughout the block.

Coefficients:	Estimate	Std. Error	z value	Pr(>|z|)	
(Intercept)	-4.062	3.403	-1.194	0.233	
Aspect	0.006	0.002	3.035	0.002	**
Rugosity	9.082	3.472	2.615	0.009	**
Bathymetry	0.072	0.013	5.599	2.16E-8	***

Summary of the probability results from the GLM performed on Area 5 to predict the probability
of finding *S. rosaceus* throughout the block.

Probability of Occurrence	Area (m^2)	Area (km^2)	Percentage of Area 5
0.0-0.1	2469951	2.470	11%
0.11-0.2	1410840	1.411	7%
0.21-0.3	1440540	1.441	7%
0.31-0.4	1937934	1.938	9%
0.41-0.5	1918836	1.919	9%
0.51-0.6	2109465	2.109	10%
0.61-0.7	2322387	2.322	11%
0.71-0.8	2806065	2.806	13%
0.81-0.9	3443193	3.443	16%
0.91-1.0	1778985	1.779	8%
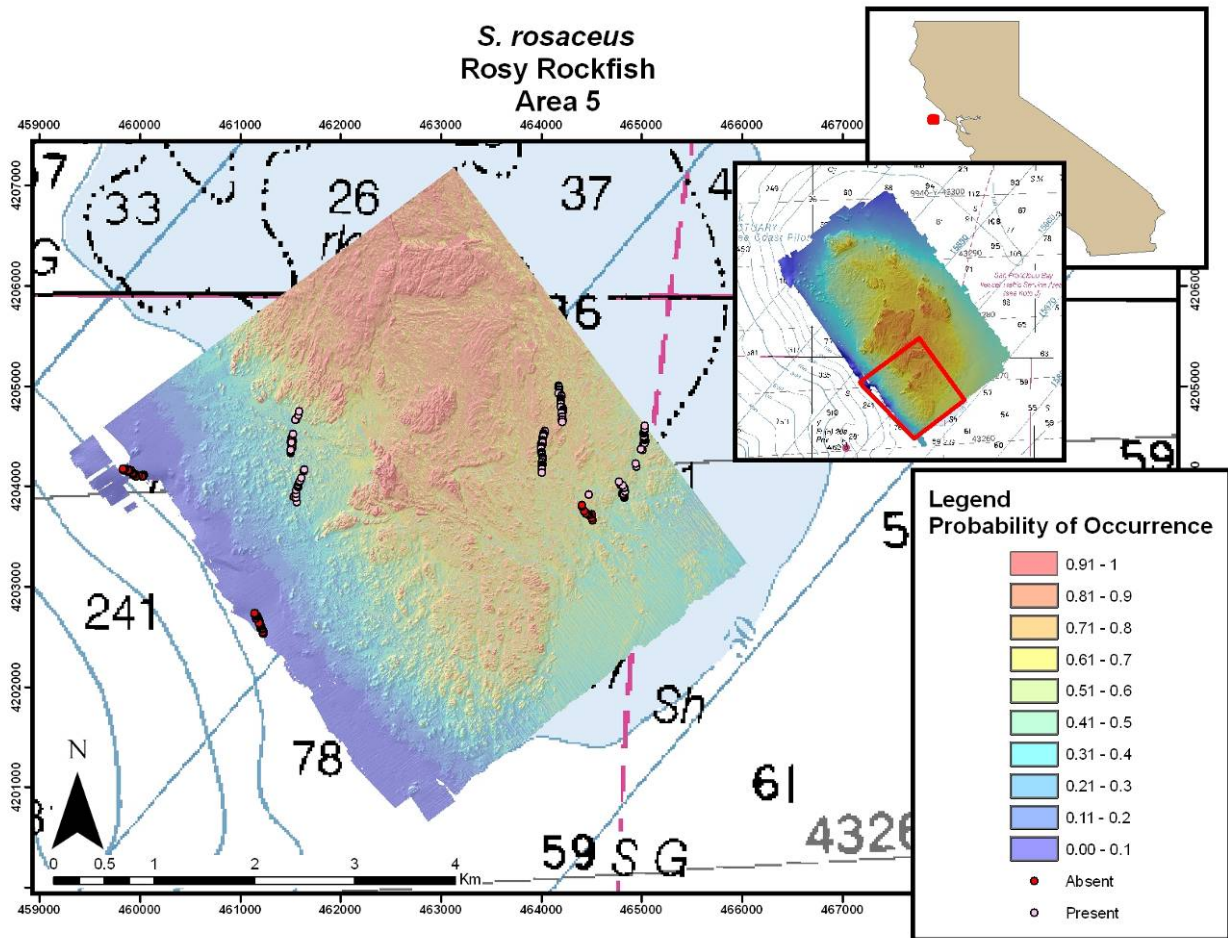
Area 5 GLM results for *Sebastes rosaceus* (rosy rockfish). Warmer colors indicate a high predicted probability of S. rosaceus occurrence, while cooler colors indicate low probabilities. Pink dots signify locations where *S. rosaceus* were observed and the red dots indicate locations where no fish were present. (Image resolution: 3m; coordinate System: UTM 10N WGS 84; NOAA Chart 18640, soundings in fathoms; multibeam survey completed fall 2005.)

AREA 6

Contingency table of the data used to determine the probability accuracy from the GLM performed in Area 5 to predict the probability of finding *Sebastes rosaceus* throughout the block (n = 119).

| Point Type | # of Points where model indicates | | Total |
	Low Probability	High Probability	
Present	8	50	58
Absent	59	2	61
Total	67	52	119

Summary of the probability results from the GLM performed on Area 6 to predict the probability of finding *S. rosaceus* throughout the block.

Coefficients:	Estimate	Std. Error	z value	Pr(>|z|)	
(Intercept)	20.439	4.061	5.032	4.85E-7	***
Substrate	4.769	2.242	2.128	0.033	*
BS Slope Position	-1.545	0.808	-1.910	0.050	*
Bathymetry	0.210	0.043	4.871	1.11E-6	***

Summary of the probability results from the GLM performed on Area 6 to predict the probability of finding *S. rosaceus* throughout the block.

Probability of Occurrence	Area (m^2)	Area (km^2)	Percentage of Area 6
0.0-0.1	12684537	12.685	63%
0.11-0.2	545472	0.545	3%
0.21-0.3	400365	0.400	2%
0.31-0.4	282276	0.282	1%
0.41-0.5	338265	0.338	2%
0.51-0.6	549720	0.550	3%
0.61-0.7	803691	0.804	4%
0.71-0.8	1132542	1.133	6%
0.81-0.9	1285353	1.285	6%
0.91-1.0	1793304	1.793	9%