Methods to incorporate known habitat in indices of abundance an applications to management

Melissa Hedges Monk[a,1,*], Rebecca R. Miller[b], E.J. Dick[a], Grant Waltz[1], Dean Wendt[c]

[a]*Southwest Fisheries Science Center, 110 McAllister Way, Santa Cruz, 95060,*

[b]*University of California Santa Cruz, Street Address, Santa Cruz, 95060,*

[c]*California Polytechnic State University, Street Address, San Luis Obispo, 93407,*

Abstract

Indices of abundance developed from fishery-dependent data are typically subject to a number of assumptions about the area and habitat fished due to the aggreation of the catch at the level of a fishing trip. In California, two surveys occur onboard the recreational charter boats fleet and samplers record location-specific data on the catch and effort during individual fishing stops throughout a trip. This location specific information coupled with high-resolution maps of the bottom substrate allowed us to subset the survey data to areas of rocky reef habitat. The six species of rockfish (*Sebastes* spp.) modeled in this paper as example all have high affinity to rocky habitat. We compared the indices of abundance developed from data filtering at the finest scale of a fishing drop using the maps of rocky reef to filter the data to the same data using only county of landing as an indicator of location. In addition, we aggregated the data across a trip to mimic data available from a dockside survey that occurs after a fishing trip to further explore the effect of datacourseness on data filtering an indices of abundance. For the data without any fishing location identified we applied the commonly used Stephens-MacCall method to identfy samples for the indices of abundance. The identification of the rocky reefs also allowed us to weight the index of abundance by the area of available habitat with predefined regions. We show that in general the

Keywords: fisheries dependent data, habitat association, groundfish, index of abundance

1. Introduction

Indices of abundance are commonly used to provide a stock assessment model with information about the stock's trend over time (Harley et al., 2001; Hilborn and Walters, 1992). For many fish stocks, only fishery-dependent survey data are available. Fishery-dependent survey data are more readily available than fishery-independent scientific survey data due to factors including the lower cost to collect data, more frequent sampling opportunities, and ability to collect data at large spatial scales where the fisheries operate.

Modelling fishery-dependent data requires making a number of assumptions due to the nature of the data being reliant on the behavior of the fishing fleet. A common metric for modelling fishery-dependent data is catch per unit effort (CPUE), which is often used under the assumption that the estimated trends are proportional to the true abundance of the stock (Maunder and Punt, 2004). However, catch rates are more likely to reflect local densities than total abundance (Haggarty and King, 2006; **?**), in which case standardized trends in CPUE (relative density) should be multiplied by habitat area, when available, to

*Corresponding author

Email addresses: melissa.monk@noaa.gov (Melissa Hedges Monk), rebecca.miller@noaa.gov (Rebecca R. Miller), edward.dick@noaa.gov (E.J. Dick), cat@example.com (Grant Waltz), cat@example.com (Dean Wendt)

[1]This is the first author footnote.

estimate trends in relative abundance. Additionally, fishery-dependent data are reliant on the behavior of the fishermen and must be standardized to account for spatial and temporal changes in fishing activity (Campbell, 2015a; Hilborn and Walters, 1992).

An analyst must alsoconsider factors such as the targeting of multiple species when developing an index of abundance. The recreational for-hire partyboat fleet may target multiple species during a trip. The target species for a recreational trip is dependent on a number of factors including weather that could limit transit to some fishing grounds, bag limit regulations, angler preference and experience, duration of the trip, and the captain's experience level. All of these factors affect the catch during a trip. For example, a recreational trip in California, USA may set out to target a particular rockfish (*Sebastes spp.*) species associated with hard substrate, but if fishing is unsuccessful or if bag limits are reached, the captain may spend a portion of the trip targeting sanddab species (*Citharichthys spp*) that inhabit areas of soft substrate.

Therefore, an analyst must determine which samples within a survey represent the effective effort directed towards the target species. The granularity of the calculation of fishing effort is dependent on the survey. A survey that interviews an angler or group of anglers at the dock or pier after the fishing trip concludes provides fishing effort at the level angler-days or angler hours. On the opposite end of the spectrum is an onboard observer survey (onboard survey) where a sampler rides along during a trip and records information on the catch and effort, often from a subset of anglers, at every fishing location during the trip. For these data, both temporal resolution (trip vs. drift) and spatial resolution (location of landing vs. location of fishing) are improved.

Here we focus on data available from fishery-dependent onboard observer surveys of California's recreational partyboat fleet, also referred to as commercial passenger fishing vessels (CPFV). The onboard observer data provide an opportunity to explore what information we gain from explicit knowledge of fishing locations. There are two surveys of the California recreational CPFV fleet (Monk et al., 2014). The California Department of Fish and Wildlife (CDFW) surveys active ports throughout the state and the California Polytechnic State University San Luis Obispo (Cal Poly) surveys vessels with home ports in San Luis Obispo County. In addition, we are able to utilize high resolution bathymetric data to define appropriate habitat for a target species.

It is not often the case where high-resolution habitat data and fishing location information are both available, and for many fishery-dependent surveys an analyst will have to determine which subset of the data to use based on available information. A widely-used method to filter fishery-dependent data to determine the samples representing effective fishing effort for the target species is the Stephens-MacCall method (2004). The Stephens-MacCall method is a binomial regression model used to predict the probability of encountering the target species in a sample, based on the presence or absence of a suite of co-occurring species. The Stephens-MacCall model was originally developed to recreational trips that fished in habitat likely to contain the target species. This method is commonly used to filter data that were collected dockside after a vessel returned to port or when location data were not provided. The availability of a 20+ year time series of onboard observer data from California

CPFVs, coupled with high-resolution habitat maps of rocky reef habitat, provides us with an opportunity to evaluate the effectiveness of the Stephens-MacCall method and compare standardized indices of abundance derived from data sets that differ in terms of spatial and temporal resolution.

One of the major recreational targets in California are bottomfish species, a group of which includes dozens of rockfish species (*Sebastes spp.*). Rockfish species' affinity to rocky habitat differs by species and ranges from the territorial gopher rockfish (*S. carnatus*) that resides within rocky crevices and maintains small home ranges of less than $20m^2$ (Larson, 1980), to a the schooling, mid-water black rockfish (*S. melanops*) that haa an average home range of $0.25km^2$, and exhibits diurnal movement offshore (**?**). Tagged black rockfish have also been recovered hundreds of miles from the initial capture location (California Collaborative Fisheries Research Program, unpublished data). The association of *Sebastes* with rocky habitat makes them ideal candidates for exploring the ability to predict effective effort from fishery-dependent data based on known habitat. In addition, the area of known rocky reef habitat data creates an opportunity to weight the index of abundance by the calculated area of rocky reef habitat along the California coast.

To explore how indices of abundance change depending on assumptions made about filtering a data set to represent the effective effort, we utilized the onboard survey data to create data sets and indices of abundance at three levels of data coarseness. At the finest scale we utilized the fishing drift level data with known location from the onboard observer surveys and subset data based on the proximity to rocky reef habitat. For this data set we also evaluated the effect of weighting by reef habitat area. We then treated the drift-level data as if the location of the individual fishing drifts were not available, and lastly, we aggregated the drift-level catches to the level of a single trip. For the last two two cases where we removed the location information, we filtered the data using the Stephens-MacCall method. We applied these methods across six nearshore rockfish species with different life histories, habitat preferences and commonness in the data.

2. Methods

We developed indices of abundance for six species or species pairs of rockfish (*Sebastes spp.*) that are of management interest on the U.S. West Coast: black rockfish (*S. melanops*), the blue and deacon rockfish (*S. mystinus*/*S. diaconus*), brown rockfish (*S. auriculatus*), China rockfish (*S. nebulosus*), gopher rockfish (*S. carnatus*), and the vermilion and sunset rockfish (*S.miniatus*/*S. crocotulus*). The two cryptic species pairs (blue/deacon and sunset/vermilion rockfish) are genetically identifiable, but not separable within the onboard observer survey time series. These species all have different latitudinal distributions, exploitation histories, and habitat and depth preferences(Love et al., 2002).

2.1. Survey Data and Habitat-based Filtering

The CDFW began a fishery-dependent onboard observer survey of the CPFV fleet in 1999. In 2004, the survey was integrated into the CDFW's California Recreational Fisheries Survey (CFRS), designed to estimate catch and effort by the recreational fleet. In response to a request from the fishing industry, Cal Poly began a supplemental onboard observer survey in

2001 of the CPFV fleet based in Port Avila and Port San Luis along the Central Coast [#fig-map]. Both the CDFW and the Cal Poly onboard observer surveys continue through present day; however, due to both spatial and temporal changes to recreational fishing regulations we limited the data for this research to the years 2004 to 2016. Between 1999 and 2003, the recreational regulations evolved from no restriction on the number of lines or hooks an angler could deploy to a one line and two-hook maximum, as well as implementation of depth restrictions. Subsequent management allowed a relaxation of depth restrictions beginning in 2017 that shifted fishing effort relative to the 2004-2016 period (Monk et al.).

While only a small portion of the total CPFV trips taken are sampled as part of the onboard observer survey, the onboard observer survey collects a large amount of data during each trip. During each trip the sampler records information for each fishing drift, defined as a period starting when the captain announces “lines down” to when the captain instructs anglers to reel their lines up. Just prior to the start of each fishing drift, the sampler selects a subset of anglers to observe, at a maximum of 15 anglers per drift. The sampler records all fish encountered (retained and discarded) by the subset of anglers as a group, i.e., catch cannot be attributed to an individual angler. Samplers also record the start and end times of a drift, location of the fishing drift (start latitude/longitude and for most drifts, end latitude/longitude), and minimum and maximum bottom depth. Fish encountered by the group of observed anglers are recorded as either retained or discarded. This provides information on the catch (count of each species) and effort (time and number of anglers fished) during each fishing drift. While both surveys include records of discarded fish, we only used the retained catch in these analyses. Discarded fish can often represent a different size structure than retained fish, either due to size limits or angler preference, or represent fish encountered during a temporal or spatial closure.

The National Marine Fishery Service's Southwest Fisheries Science Center (SWFSC) developed a relational database for the CDFW onboard survey (2014) that is updated annually. The Cal Poly data are also provided to the SWFSC annually. All data were checked for potential errors at the drift-level by SWFSC staff. The data sets were filtered for errors within the relational database before analyses were conducted, and the data used here reflect changes from the QA/QC process that may not be reflected in the raw data available directly from the CDFW.

The CPFV data in this paper included only areas north of Point Conception ($34°27'N$) due to gaps in habitat coverage further south. Point Conception is a significant biogeographic boundary (Valentine (1966)), and the composition of the fish communities in southern California differ, potentially reducing the effectiveness of methods that rely on species associations, such as the method of Stephens and MacCall. The recreational fisheries south of Point Conception are also fundamentally different, with a higher percentage of trips targeting mixed species and pelagic and highly migratory species, as well as less rocky habitat in nearshore areas compared to northern California.

We removed drifts in the upper and lower 1% of the recorded time fished and recorded observed anglers, as these may not accurately define a successful fishing drift or may represent data errors. Similarly, we filtered the data to retain 99% of all drifts based on average drift depth. We calculated average depth from the recorded minimum and maximum depths when available or the imputed minimum and maximum depth from the bathymetry layer described below.

High resolution seafloor mapping data of the study area allowed us to overlay the starting latitude/longitude of each drift from the onboard observer surveys with predicted habitat (referred throughout the paper as the drift-level, habitat-informed data). Specifically, we used bathymetry and backscatter data collected by the California Seafloor Mapping Program (CSMP) (Golden, 2013; Bay, 2014). The CSMP mapped California state waters at a 2 m resolution from Point Conception to the California-Oregon border. We created a mosaic from 137 CSMP substrate blocks that ranged in size from 16 km^2 to more than 400 km^2. The CSMP identifies rough and smooth substrates, surface:planar area, and a vector ruggedness measure (VRM) of the bathymetric digital elevation model [#fig-map2]. The CSMP set a varying VRM threshold for each of the substrate blocks, removed any artifacts, and their product is considered a conservative estimate of rough habitat.

The digital mosaic of137 CSMP substrate raster blocks with pixels designated as rough habitat (our rocky habitat proxy) was then converted from a raster format to polygons. Next, a 5 m buffer region was created around the rough habitat polygon to allow for any small errors in positional accuracy. The area of each reef polygon was calculated, and those reefs greater than or equal to $100m^2$ were included. All spatial analyses were conducted using ArcMap 10.3 (ESRI citation).

Contiguous polygons defined in this way as rocky reef substrate were treated as a single rocky reef, regardless of size. Polygons with at least 200 m distance between them were treated as separate reefs. The area of rocky habitat for this paper includes reefs within California state waters (up to 3 nm from shore). The mapped area does not include very shallow areas close to shore, which extend approximately 200-500 m from the shoreline, due to gaps in habitat mapping data However, fishing by the CPFV fleet is also limited in these waters due to shallow depths and the presence of thick kelp beds. We assigned fishing drifts to reefs based on the recorded start location of a drift, given that the end locations were not always recorded.

For each target species, we calculated the cumulative distribution of distance to rocky reef (in meters) for drifts that retained the target species and used a distance cutoff of 90% for each species. To illustrate the similarities and differences among the six species, we mapped he percentage of fishing drifts within an aggregated region where each species was present and retained. To illustrate differences in relative densities among species we calculated the average CPUE, before standardization, for each species and aggregated area. We also examined the assumption that effort might be proportional to reef area by obtaining effort estimates for the CPFV trips (cite RecFIN) and comparing the area of rocky habitat to the effort in each region .

2.2. Stephens-MacCall Data Filtering

To identify effective fishing effort without the use of precise fishing locations and habitat data, we began our analyses again from the unfiltered onboard observer data set. In this way, data filtered by the Stephens-MacCall method would solely rely on information about species composition of the catch, as would be the case in the absence of habitat information. Since the Stephens-MacCall method relies on species composition of the catch, one can imagine that aggregating catch across fishing locations could affect which records are retained for CPUE standardization. To illustrate the impact of having less spatially-explicit data on both data filtering and the resulting indices of abundance, we applied the Stephens-MacCall method using two levels of aggregation (catch by drift and by trip) (2004). For the Stephens-MacCall drift-level data we removed all location and depth identifiers for a drift and kept the county of landing as a spatial identifier. To construct a data set that mimicked trip-level data, we took the drift-level data, aggregated the observed retained catch within a trip, and kept the county of landing as a spatial identifier.

Before applying the Stephens-MacCall method, we identified a suite of potentially informative predictor species for each of the six target species. Species present in fewer than 1% of all drifts and 3% of all trips were removed to reduce the number of species to those that were informative. A lower threshold of 1% was selected for the drift-level data due to the change in magnitude of the number of samples when using drifts versus trips.

The goal when implementing the Stephens-MacCall method is to eliminate trips with a low probability of catching the target species given the other species caught on the trip. Stephens and MacCall proposed excluding samples from index standardization based on a criterion of balancing the number of false positives and false negatives from the predicted probability of encounter. The false positives are samples that are predicted to encounter the target species based on the species composition of the catch, but did not. The false negatives are samples that were not predicted to encounter the target species, given the catch composition, but caught at least one target species. Balancing the false negatives and false positives remains common practice and is the method we applied in this study.

2.3. Indices of Abundance

We generated four standardized indices of abundance for each of the six species; one index for each of the three data filtering methods (drift-level habitat-informed, drift-level Stephens-MacCall, trip-level Stephens-MacCall), plus an area-weighted index from the habitat-informed drift-level data. All indices were modeled using Bayesian generalized linear models (GLMs) and the delta GLM method (Lo et al., 1992; Stefánsson, 1996). The delta GLM method is commonly used to standardize catch-per-unit effort for stock assessments [citations]. The delta method models the data with two separate GLMs; one for the probability of encountering the species of interest using a binomial likelihood and a logit link function, and the second GLM for the positive encounters assuming either a gamma or lognormal error structure. The error structure of the positive model was selected via the Akaike Information Criterion (AIC) from models with the full suite of considered explanatory variables.

The response variable for the positive models was angler-retained catch per unit effort. For the indices modeled at the level of a drift, effort was calculated as the number of angler hours fished on a drift. The trip-level effort was calculated as angler days, using the average number of observed anglers across all drifts on a trip.

To keep comparisons across data filtering methods similar, depth was not considered as an explanatory variable in the habitat-informed index. Depth is often a significant explanatory variable for rockfish species, with many rockfish species and populations separated by depth (Love et al., 2002). Year was always included in as an explanatory variable in model selection, even if it was not significant, because the goal of the index of abundance was to extract the year effect. Other explanatory variables considered for the habitat-informed index were aggregated regions of rocky reef (categorical variable) and wave (a 3-month aggregated period of time, e.g., January-March). The area-weighted index also included a year/rocky reef interaction term, even if it was not statistically significant, to allow us to weight the index by the area of rocky reef. The regions of rocky reef were aggregated differently for each species to ensure adequate sample sizes to explore the year/rocky reef interaction.

Explanatory variables for the two indices based on data filtered using the Stephens-MacCall method (blind to habitat information at the drift- and trip-level) included only year, wave and aggregated counties of landing. California has 14 coastal counties north of Point Conception, 11 of which were represented in these data. We aggregated the northern counties of Del Norte, Humboldt and Mendocino into one region, Sonoma and Marin counties just north of San Francisco into another region and Alameda and San Francisco counties into a third region. The remaining counties of San Mateo, Santa Cruz, Monterey and San Luis Obispo were not aggregated.

Model selection for the binomial and positive observation models was based on AIC using the lme4 package in R, and unless very different predictors were selected, the same predictors were used in each of the two Bayesian models. The Bayesian models were run with 5,000 iterations and weakly informative priors. Posterior predictive model checks were examined for both the binomial and positive observation models, including the predicted percent positive compared to the maximum likelihood estimates. We constructed the unweighted annual index by multiplying the back-transformed posterior draws from the year coefficients from the binomial model by the exponentiated positive model draws from the year coefficients, and taking the mean and standard deviation of the distribution of the product for each year.

The area-weighted, habitat-informed index was developed by extracting the posterior draws of the unweighted index, and then summing across the product of the back-transformed posteriors weighted by the fraction of total area within each reef. To compare the indices across the three data filtering methods and the area-weighted index, each index was scaled to its mean value.

3. Results

3.1. Survey Data and Habitat-based Filtering

Prior to any filtering a total of 19,425 drifts that aggregated to 2,270 trips were available for the analyses. The number of initial samples used for the Stephens-MacCall filtering method were higher than the habitat-informed data described in the previous section because retained drifts with missing locations (latitude/longitude). Approximately 21% of all the CPFV trips observed by CDFW from 2004-2016 occurred north of Point Conception. It is important to note that north of Bodega Bay, California, the majority of charter boats are smaller six-pack vessels that may not have the capacity to carry a sampler onboard. As a result, sample sizes in this part of the state are smaller than areas to the south. The addition of the Cal Poly onboard observer survey to the CDFW survey more than doubled the sample sizes of observed trips in San Luis Obispo County, with an average annual increase of 155% from 2004-2016.

From 2004-2016 the drift-level data contained a total of 19,425 fishing drifts, and after removing drifts with missing effort information (time fished and/or observed anglers), 19,180 drifts remained. The filter removing the upper and lower 1% of the time fished and number of observed anglers resulted in fishing drifts lasting between three and 96 minutes and three to 15 observed anglers, and reduced the data to 18,591 fishing drifts. The remaining data filter for depth resulted in a cutoff of 46.6 fathoms, and retained 18,405 drifts based on average drift depth. A filter on the minimum depth was not included here because the recreational fleet was not limited to a minimum fishing depth and all of the fishing drift locations were verified during the QA/QC process. In the final, filtered drift-level data set the average time fished was X minutes with a standard deviation (SD) of X. The average number of observed anglers was X (SD=X), and average estimated depth was X (SD=X).

We define 108 areas of rocky habitat within California state waters from the California/Oregon border to Point Conception. The two meter resolution of the substrate shows the patchiness and heterogeneity of the rocky substrate (Figure 1). We characterize rocky habitat using thresholds as determined by the CSMP (Bay (2014)). While the location-specific data from the fishing fleet is governed by confidentiality and cannot be displayed here, 85% of the fishing drifts were within 5 m of rocky habitat. The recreational fishing fleet's targeting of rockfish species was verified by the distributions of the distance from rocky habitat for each of the six species. The distance from rocky habitat cutoff (retaining 90% of drifts encountering each species) was six meters for blue, China and gopher rockfish, eight meters for vermilion rockfish, 14 meters for black rockfish and 16 meters for brown rockfish. The percentage of drifts and trips encountering the target species can be found in Table 2.

Based on exploratory analyses and consideration of the available data,we aggregated the areas of rocky habitat grouped into six regions to ensure adequate sample sizes for developing indices of abundance (Figure 2). While covering a small area (5% of the rocky habitat), the number of observed fishing drifts within state waters around the Farallon Islands off the coast of San Francisco was high enough to warrant keeping it as a separate area of rocky habitat. The region defined from the California/Oregon border to San Francisco encompasses 49% of the total rocky habitat in state waters by area, but only 12% of the observed drifts (2,637). Each of the four remaining regions of rocky habitat defined from San Francisco to Point Conception contained an average of 12% of the available habitat (Table X). The CDFW estimated fishing effort by management district, which does not exactly align with our areas of grouped reef habitat. Only considering the fishing effort north of Point Conception, CDFW estimated an average of 9% of the CPFV trips occurred from the California/Oregon border through Mendocino County, 38% from Sonoma through San Mateo County, and 53% from Santa Cruz to Point Conception.

The differences in latitudinal distribution of the six species is apparent from the maps of percent of positive observations (Figure 3). The distribution of black rockfish tapers off south of San Francisco, whereas percent of fishing drifts encountering vermilion, gopher, and blue rockfish are higher south of San Francisco. Brown rockfish is distributed across all of coastal California, with slightly higher encounter rates south of San Francisco and the percentage of drifts retaining China rockfish was low coastwide. The average CPUE was highest for blue rockfish between San Francisco south to Big Sur (Figure 4). The average CPUE for black rockfish average was higher in the north, while gopher rockfish CPUE was generally consistent across the coast, albeit slightly higher south of Big Sur. China rockfish CPUE catch was typically low coastwide, with slightly higher catch rates in the Farallon Island reefs.

The final aggregation of the reefs and total area within each region are found in Table 1. The fraction of drifts retained for the indices of abundance was high for all six species (80% or greater), indicating that fishing effort represented by these data occurred mainly near areas of rocky habitat.

3.2. Stephens-MacCall Data Filtering

A total of 19,425 drifts that aggregated to 2,252 trips were used for the trip-level Stephens-MacCall filtering. In general, the co-occurring species used for the Stephens-MacCall method were similar for the drift-level and the trip-level data. We present the coefficients and from the binomial GLM with 95% confidence intervals for black rockfish at the trip-level and drift-level in Figure 5. The corresponding plots for the remaining species are in the supplemental material. The confidence intervals were consistently larger for the trip-level data and the co-occurring species at the drift-level provide a refined look at species that have positive coefficients. For black rockfish, a noticeable difference is the intercept. At the trip-level the intercept (probability of catching the target species, given that none of the indicator species were caught) is uninformative and at the drift-level the intercept is strongly negative. A higher fraction of the co-occurring species are uninformative information about the target species in our study (i.e., 95% confidence interval crosses zero) for the trip-level data than the drift-level.

The remaining species all co-occurred with the target species in at least one trip and were retained for the Stephens-MacCall logistic regression. Coefficients from the Stephens-MacCall analysis (a binomial generalized linear model) were positive for species that are more likely to co-occur with the target species, and negative for species that were less likely to be caught with target species. The intercept represented the probability of observing only the target species in a sample.

The percentage of samples retained for each data filtering method differed by species, but followed the general trend that the lowest percent of samples were retained from the Stephens-MacCall filtering at the drift level, ranging from 12% of samples retained for China rockfish and 54% for blue rockfish (Table 2). A higher percent of samples were retained both from the other two methods, with an average of 83% of drifts retained when habitat was included as a filter. Data filtering for the trip-level indices that retained all positive observations resulted in a high proportion of positive samples (0.70 - 0.86) for all species.

To determine how consistent the Stephens-MacCall trip-level filter was with the habitat-informed filter, we looked at the distance to reef from all of the drifts contributing to trips that were used for the trip-level index. This provides a proxy for how well the Stephens-MacCall method infers habitat from species associations. Using the same distance from reef cutoff by species as calculated from the habitat-informed data, we calculated the percentage of drifts that were further from a reef than would be expected, but used in the data to develop the trip-level index. The percentage of drifts contributing to the trips outside reef habitat was 11% for black rockfish, 13% for blue rockfish, 10% for brown rockfish, 12% for China rockfish, 12% for gopher rockfish, and 11% for vermilion rockfish.

3.3. Indices of Abundance

Model selection using AIC resulted in all but three of the 24 indices of relative abundance bineg modeled with a lognormal distribution for positive observations. The trip-level indices for black, blue and gopher rockfish were modeled using a gamma (refer to the supplementary material for AIC scores). All of the covariates (year, reef, and wave) were selected as main effects for both the binomial and positive models for all species in the habitat-informed drift level index. For instances where the wave. However, the difference in AIC relative to the best model (delta-AIC) was less than ten so we chose to maintain the model with year, reef and wave.

(LMK and I can put these in a table in the doc) The full model that included the reef:year interaction was selected by AIC for all species except for China rockfish. For China rockfish the positive binomial model selected the interaction covariate, but the model without the interaction was select for the positive lognormal model by an difference in AIC of 22.

However, in order to look at the effects of the area-weighting on the index, we included the year:reef interaction in the final model for China rockfish.

For both the drift-level and trip-level Stephens-MacCall filtered data, year, county and wave were selected for black rockfish, blue rockfish, gopher rockfish, and vermilion rockfish and the drift-level index for brown rockfish. The model incorporate in year and county was selected for the trip-level Stephens-MacCall filtered index for brown rockfish and both Stephens-MacCall filtered indices for China rockfish.

In general, the larger increases and decreases in the indices were similar among the four indices developed for each species (Figure 6). The generalized approach used in this paper to create indices with comparable methods resulted in different results for each species. The area-weighted indices are reflective of the total available habitat and use all of the available high resolution habitat and fishing drift data. However, differences among the four indices were different for each species. The average CVs between the drift-level area-weighted index and the drift-level habitat informed indices were similar, as expected, since they both used the same data with the only difference being the year:area interaction in the models (Table 3). However, the average CV between drift-level habitat-informed filtering and Stephens-MacCall filtering for the drift-level data differed by species.

The area-weighting for black rockfish, a species distributed predominantly north of Santa Cruz, California did have an effect on the index for a number of years, most notably in 2013 where the area-weighted estimate is lower than all three other indices(Figure 6a). The effect of the area-weighting is also apparent for black rockfish in 2005, 2007, and 2009.The average CV decreased from the trip-level index (0.671) to to the area-weighted index (0.443). Interestingly, the average CV was lowest overall for the drift-level Stephens-MacCall index (0.364) which also modeled much smaller data with a high proportion of positive catches of black rockfish (Table 2).

Blue rockfish is ubiquitous across the study area and was one of the two species for which the index was weighted by the six regions of rocky reef habitat. The area-weighted index differs from the other three in 2006 with an estimated higher relative abundance and in 2014 with an estimated lower relative abundance. Even during the years from 2009 to 2012 when the estimated relative abundance was low for all of the indices, there were differences among the four trends with the drift-level habitat-informed index estimating the lowest relative abundance.

All four indices for brown rockfish suggested differing trends, with this species having the highest estimated error for both the trip-level and drift-level Stephens-MacCall filtered data (Figure 6c). In ten of the years the area-weighted index estimated a either the largest or smallest relative abundance compared to the other indices. For brown rockfish the two habitat-informed indices were more similar than the Stephens-MacCall filtered data. The average CV for brown rockfish from the Stephens-MacCall filtering was large (0.679) compared to the habitat informed filtering (0.142).

China rockfish is the only species for which the trip-level Stephens-MacCall filtered index had the lowest average coefficient of variation that increased with the habitat-informed filtering (Table 3). Although the trends among the four indices were similar, this is the only species for which the highest error was consistently estimated for both habitat-informed drift-level indices (Figure 6d). China rockfish is one of the less common species observed in the data with the highest average CPUE from catches the Farallon Islands, which is an overall small percentage of the total habitat (Table 1).

The observed trends for gopher rockfish were similar among all indices and the trip-level Stephens-MacCall index had the highest average CV (0.626) compared to the average CVs of less than two from all of the other drift-level indices. China rockfish is the only species for which the trip-level index had the lowest average coefficient of variation, which increased with the habitat-informed filtering . For all other species, the habitat-informed filtering resulted indices with a lower average CV than the trip-level filtering.

The indices of relative abundance for vermilion rockfish were relatively similar in trends across the time series (Figure 6f). Vermilion rockfish is the second species for which all six areas of rocky reef habitat remained dis-aggregated in the models. For vermilion rockfish, while the trends are similar among all four indices, the effect of area-weighting dampens the increase modeled from the habitat-informed drift level data from 2004-2006, where the area-weighting down-weighted the relative abundance from the drift-level habitat informed index.

4. Discussion

Fishery-dependent indices of abundance will continue to be incorporated in fisheries stock assessments. We demonstrated the effects of subsetting fishery-dependent survey data to samples representing effective effort at varying levels of data resolution. The estimated indices of abundance illustrated the changes in trends and variance to create a subset of samples representing the effective effort for a target species, and how that selection affected the trends in indices of relative abundance. The combination of fine-scale CPUE data coupled with the available habitat data creates allows us to model an index of relative density, rather than abundance. The fishery-dependent onboard observer survey conducted by CDFW and Cal Poly is a benchmark recreational fishery-dependent time series. The survey provides many elements that would usually only be collected from research surveys, including fishing locations, fishing depth, time fished, and speciated catch and discard information, and currently has over a 20 year time series.

We also demonstrated that the habitat-informed data filtering provides a method to select samples with effective fishing reduces the subject decision points required when filtering multispecies data by utilizing known habitat characteristics. This also allows us to create an area-weighted index that accounts for variable species density along the coast. This not only addressed a key assumption of identifying effective fishing effort for a multispecies fishery, but also appropriately weights the sample data based on the known area of habitat.

The Stephens-MacCall filtering method has several subjective decision points, including which species to include in the analysis, the threshold to determine which samples to retain or remove, and the spatial extent of data to include. The Stephens-MacCall filter is useful in identifying co-occurring or non-occurring species, but it assumes all effort was exerted in pursuit of a single target species. Stephens-MacCall filtering is most often used for data collected at the trip-level in the absence of known fishing locations. If more than one species or species complex was targeted during a trip it can result in co-occurrence of species in the trip-level catch that do not truly co-occur. This was clearly shown in the differences between the trip-level Stephens-MacCall filtering and the drift-level Stephens-MacCall filtering that reflects species co-occurrence at a finer scale. If the fishing drifts covered small enough areas the Stephens-MacCall filter at the drift-level inherently contains information on habitat preferences and community structure.

The choice of a threshold value to use from the Stephen-MacCall method has been a topic explored within stock assessments for both commercial and recreational data (Dettloff (2021); Cope et al. (2015);Ducharme-Barth et al. (2018)). There is currently no guidance on best practices for the decision points in the Stephens-MacCall method that may lead to additional bias in data selection. For instance, all of the observations in the onboard observer survey are recorded by trained samplers who are assumed to correctly identify species. With this assumption, we retained all of the samples observing the target species regardless of the probability estimated from the Stephens-MacCall model. The drift-level habitat informed data retained a larger number of drifts than the drift-level Stephens-MacCall filtered data, as a result of the majority of drifts occurring over hard bottom habitat. However, one caveat of the rocky reef habitat data is that there is currenlty only a binomial classification of hard and soft substrate available, and we assume that all rocky habitat is suitable habitat. We know from the variability in rugosity and relief displayed in Figure (Figure 1) that these characteristics can change at small spatial scales. The *Sebastes spp.* complex north of Point

Conception have differential hard bottom preferences, which have been verified by visual surveys (Laidig et al. (2009); Anderson and Yoklavich (2007); Haggarty and King (2006)) and from discussions with experienced fishermen.

Based on the current practice of retaining the false positives within the Stephens-MacCall method as described in the methods section, the trip-level data are prone to overestimate fishing effort for the less common species, and result in larger variances in the indices of abundance. Looking at the number of trips selected between the drift-level Stephens-MacCall filter and the habitat-informed filter, the Stephens-MacCall filter (based on the retention of the false negatives) may exclude too many samples that fished in the appropriate habitat, but did not meet the probability threshold (Table 2). Looking at the number of trips selected between the drift-level Stephens-MacCall filter and the habitat-informed filter, the Stephens-MacCall filter (based on the retention of the false negatives) may exclude too many samples that fished in the appropriate habitat, but did not meet the probability threshold (Table 2). The Stephens-MacCall filter may be over-selecting samples where the species was not observed if the target species is less common, e.g., China rockfish, but has a strong positive co-occurrence with a more common midwater, schooling species, e.g., blue rockfish. China rockfish in particular have a heterogenous distribution with an affinity to high relief habitat (Love et al. (2002)). The Stephens-MacCall filter may be over-selecting samples where the species was not observed if the target species is less common, e.g., China rockfish, but has a strong positive co-occurrence with a more ubiquitous species, e.g., blue rockfish. For a ubiquitous species like vermilion rockfish, the Stephens-MacCall drift-level data included 51% fewer drifts than the habitat-informed data, and for the less common China rockfish, 84% fewer total samples. The Stephens-MacCall method applied at the drift-level provides insight into the fine-scale species associations, but may also reflect targetting of species that are more common or schooling. The integration of the habitat data with the onboard observer fishing drift locations provides the most accurate information for filtering the survey data. The differences between the drift-level Stephens-MacCall filtered data and the habitat-informed filter illustrate what may represent the habitat preference of individual species.

Areas of rocky habitat that were well fished and never observed the target species should be investigated to determine if the appropriate habitat exists in that area, or if other factors such as historical fishing pressure explain the lack of target species catch. The suite of six species that we modeled in this paper is a concrete example of why habitat is important and also varies among the species. The high proportion of retained drifts across species when using habitat as a data filter indicates that hate majority of drifts occurred over, or very close to, rocky habitat. Both blue and black rockfish have high affinity to rocky habitat, but occur higher off the bottom and are both schooling species. It is not uncommon to have a fishing drift dominated by blue rockfish in central California, or black rockfish further north. However, the Stephen-MacCall approach does not account for this by modeling presence/absence. Additional factors such as latitude could be included in the logistic regression to inform the Stephens-MacCall model.

The majority of groundfish species targeted by the CPFV fleet north of Point Conception during the time period of this study all have high associations to rocky habitat. In this case, the Stephens-MacCall method can be considered a proxy for habitat when the species of interest has known associations. This can be expanded in areas where trips are known to target species of interest, but no habitat data are available the proportion of trips encountering the target species could be used as a proxy for habitat.

The differences observed in the indices of abundance and knowledge of species-specific habitat preference will allow us to fine-tune these indices on a species-specific basis. The characteristics and classification of the rocky reef habitat into more specific substrate types, e.g., boulder vs pinnacle, are currently only available for a small fraction of the mapped area. Therefore, all areas of rocky substrate are currently created equal. A number of video surveys have shown habitat associations differ by species (Love et al., 2002), and the weights applied as available habitat may vary by species and be lower than the weights used in this paper. Although we did not exclude data based on the species' distributions from the indices developed here, the habitat-informed filters also allow an analyst to subset the data and exclude areas of rocky reef habitat outside of the species' range. For instance, black rockfish have been observed as far south as Point Conception, but their distribution tapers off south of Santa Cruz, California.

Fishery-dependent indices of abundance undergo higher levels of scrutiny during stock assessment reviews due to the nature of the data being driven by angler behavior. Catch from the recreational CPFV fishery is dependent on a number of factors including weather, distance from port, the clientele preferences, angler experience and captain's knowledge.

These models also do not account for distance to the nearest port, which has been shown to significantly impact the access to fish as well as historical fishing pressure (Miller et al., 2014). There are additional key assumptions made when using the onboard observer data in a stock assessment, including that fishing behavior remains the same when samplers are not onboard the vessel.

Catch from the recreational CPFV fishery is dependent on a number of factors including weather, distance from port, the clientele preferences, angler experience and captain's knowledge. These models also do not account for distance to the nearest port, which has been shown to significantly impact the access to fish as well as historical fishing pressure. Recent studies have identified the need to investigate the assumptions and uncertainty in relative indices of abundance from visual surveys (Bacheler et al., 2015; Campbell, 2015b) and simulation studies (Siegfried et al., 2016), and the same holds true for fishery-dependent surveys like the onboard observer survey. To address the potential bias in angling data for groundfish species, Haggarty and King (2006) conducted a SCUBA dive survey followed by a research angling survey directly above the dive plots and found a strictly proportional relationship between the density estimated from the SCUBA survey and CPUE from the angling survey for copper rockfish, a species whose habitat and depth distribution were well covered by the survey. Further analyses are underway to explore the fine-scale habitat characteristics to fine-tune the habitat informed data selection methods. We also plan to explore changes in fishing behavior related to management measures and and fisher behavior to explain shifts among target species or how large recruitment events for one species may affect the index of abundance for another species. While not all of these factors can be controlled for, defining the samples with effective effort will provide the most accurate index and appropriate variance for stock assessments.

removed: However, they found no relationship between SCUBA dive survey data and the angling survey for kelp greenling (*Hexagrammos decagrammus*), which the authors hypothesized was due to the greenling's avoidance of the bait used.

removed: Further analyses are underway to explore the fine-scale habitat characteristics that will allow the methods described in this paper to be fine-tuned. We also plan to explore changes in fishing behavior related to management measures and and fisher behavior to explain shifts among target species or how large recruitment events for one species may affect the index of abundance for another species.
removed: This does not hold for areas where multiple species complexes are targeted on same trip, e.g, a multi-day trip may target large pelagic species and once trip limits are reached, the trip may focus on a secondary target, which is the case for the California CPFV fleet fishing south of Point Conception.

removed: An additional source of bias in fishery-dependent data is the change in regulation over time. These can be bag limits, minimum size restrictions, and area closures that the change of available habitat. Depth restrictions have also been in place for the recreational fleet since the early 2000s, which were relaxed in 2017 and was the reason we constrained the years modeled for this study.

removed: Versions of the drift-level habitat-informed indices were approved by the Pacific

Fisheries Management Council's Science and Statistical Committee for use in the 2013 stock assessments and have been used in the stock assessment process since. Comparisons should not be drawn between the indices presented here and the stock assessment documents as the indices in this paper were simplified to develop direct comparisons among methods. When filtering and modeling the onboard observer data for a stock assessment, additional filtering steps would be taken, such as excluding areas where species are rare, e.g., south of Santa Cruz for black rockfish, inclusion of depth as a covariate in the index of abundance, and an exploration of alternative error distributions.

removed: Another example is closures and retraction of the available habitat open to fishing. California developed a network of Marine Protected Areas (MPAs) in 2007, that reduced the available rocky reef habitat to the recreational fleet by approximately 23% in state waters north of Point Conception.

5. Acknowledgements

We thank the following reviewers for comments that improved the manuscript. CDFW for collection of the onboard observer data Cal Poly for the collection of the supplemental data

Data attribution: CDFW acquires data from its own fisheries management activities and from mandatory reporting requirements on the commercial and recreational fishery pursuant to the Fish and Game Code and the California Code of Regulations. These data are constantly being updated, and data sets are constantly modified. CDFW may provide data upon request, but, unless otherwise stated, does not endorse any particular analytical

methods, interpretations, or conclusions based upon the data it provides.

6. Tables

Table 1: Area of rocky habitat in state waters aggregated to the levels modelled for each species. The merged cells for each species indicate which areas of rocky habitat were aggregated to ensure appropriate samples sizes to explore an area-weighted index.

<table>
<tr><th>Rocky Reef Desginations</th><th>Blue rockfish & Vermilion rockfish</th><th>Black rockfish</th><th>Brown rockfish</th><th>China rockfish</th><th>Gopher rockfish</th></tr>
<tr><td>California border to San Francisco</td><td>439.546</td><td>439.546</td><td>439.546</td><td rowspan="2">547.970</td><td rowspan="4">735.825</td></tr>
<tr><td>San Francisco to Santa Cruz</td><td>108.424</td><td>108.424</td><td rowspan="2">498.967</td></tr>
<tr><td>Farallon Islands</td><td>50.252</td><td rowspan="4">390.543</td><td>50.252</td></tr>
<tr><td>Moss Landing to Big Sur</td><td>137.603</td><td rowspan="2">228.027</td><td>137.603</td></tr>
<tr><td>Big Sur to Morro Bay</td><td>90.424</td><td rowspan="2">202.688</td><td>90.424</td></tr>
<tr><td>Morro Bay to Point Conception</td><td>112.264</td><td>112.264</td><td>112.264</td></tr>
</table>

Table 2: The number of samples retained after filtering to create the index of abundance with the percent of samples that caught the species in parentheses.

	Drift-level		Trip-level
Species	Habitat-informed	Stephens-MacCall filtered	Stephens-MacCall filtered
Black Rockfish	16306 (16%)	3038 (30%)	706 (68%)
Blue Rockfish	15283 (44%)	7490 (60%)	1813 (91%)
Brown Rockfish	15736 (16%)	2740 (31%)	806 (62%)
China Rockfish	14865 (8%)	1331 (22%)	798 (57%)
Gopher Rockfish	14476 (31%)	5088 (45%)	1449 (81%)
Vermilion Rockfish	14713 (30%)	5040 (45%)	1627 (85%)

Table 3: The average Coefficient of Variation (CV) for each index of abundance, where SM-filtered is the Stephens-MacCall filtering.

	Drift-level			Trip-level
Species	Area-weighted	Habitat-informed	Stephens-MacCall filtered	Stephens-MacCall filtered
Black rockfish	0.4426091	0.4493133	0.7641099	0.8495448
Blue rockfish	0.1343866	0.1415416	0.1610735	0.3324914
Brown rockfish	0.2415686	0.2399299	0.8652880	0.9161881
China rockfish	0.3196653	0.3011640	0.4481187	0.2087114
Gopher rockfish	0.1785421	0.1831132	0.2562205	0.2535190
Vermilion rockfish	0.1519120	0.1781884	0.4224451	0.5087889

7. Figures

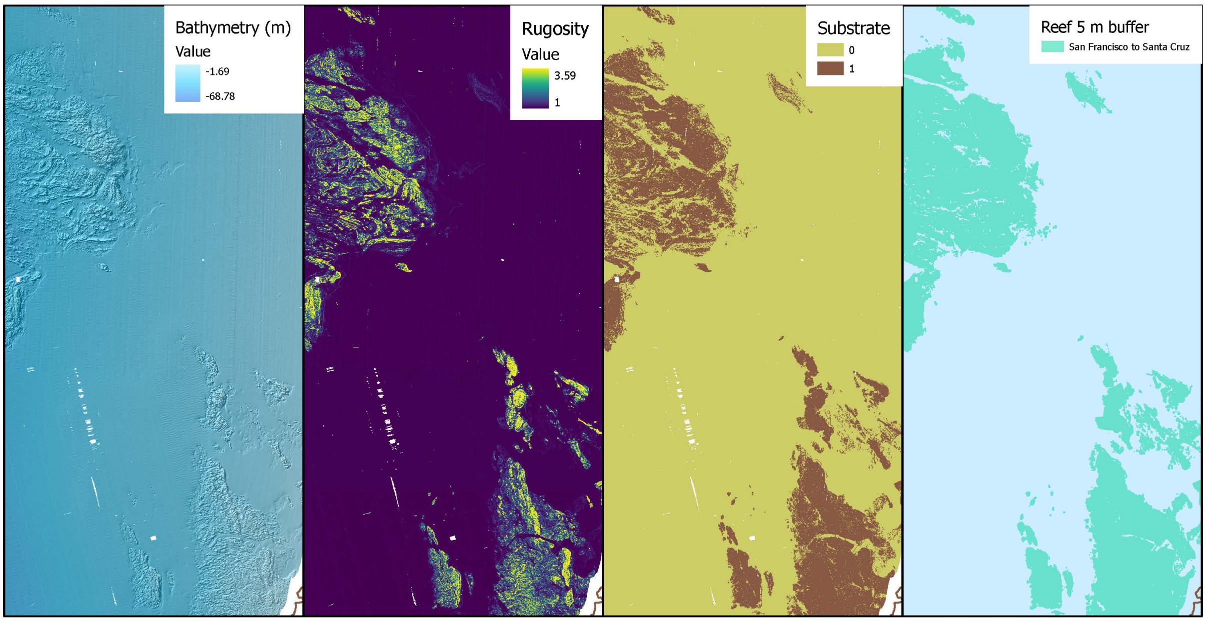


Figure 1: A example of the high resolution bathymetric data and components of bathymetry and rugosity used to define rough versus smooth substrate (where hard substrate is denoted by 1). The far right panel displays the hard substrate with the added 5 m buffer to represent the rocky reef habitat.

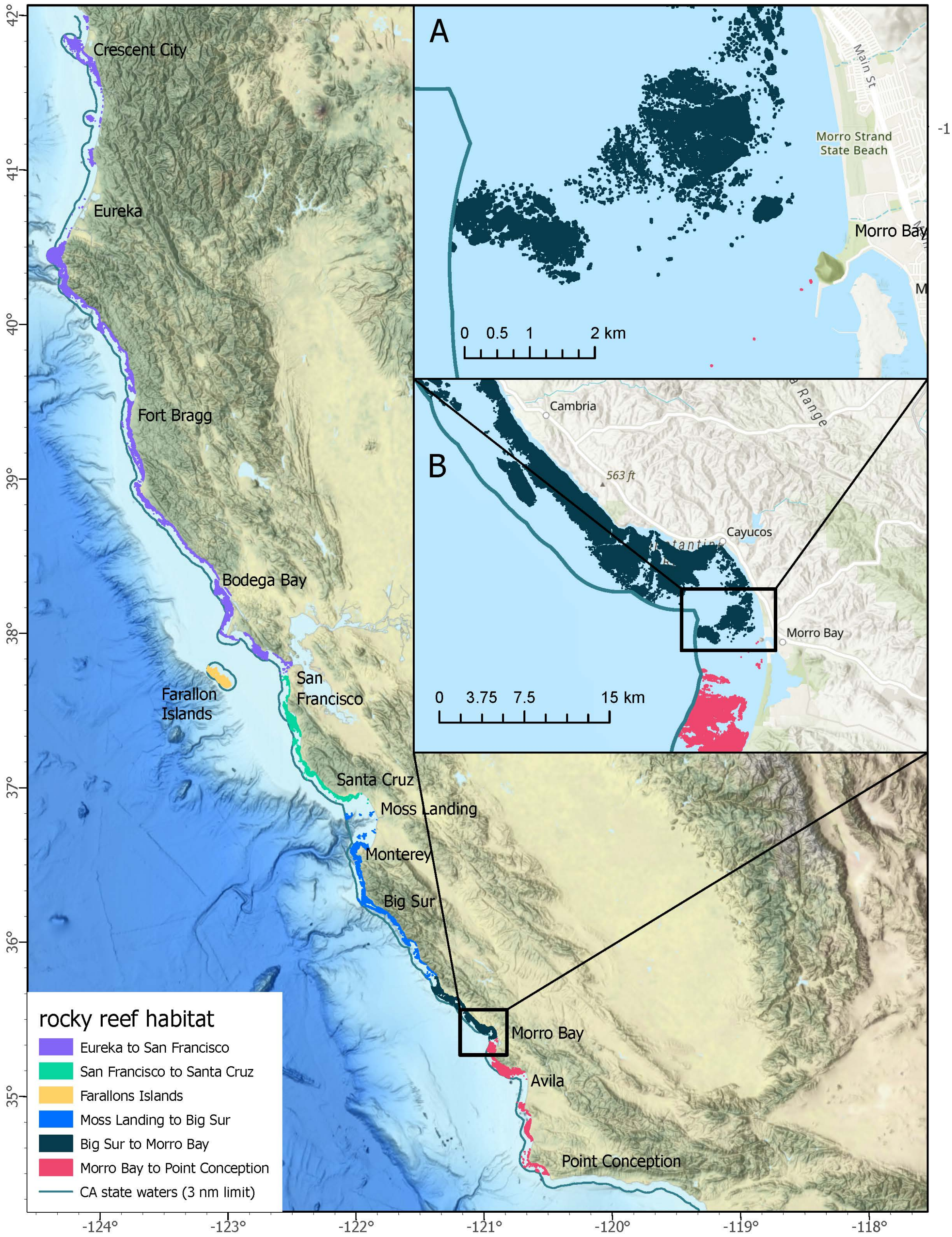


Figure 2: A maps of California state waters north of Point Conception colored by the aggregated areas of rocky reef habitat, including inset A depicting the rocky reef habitat in relation to 3 nm state water boundary state waters and inset B showing the high resolution rocky habitat in the area.

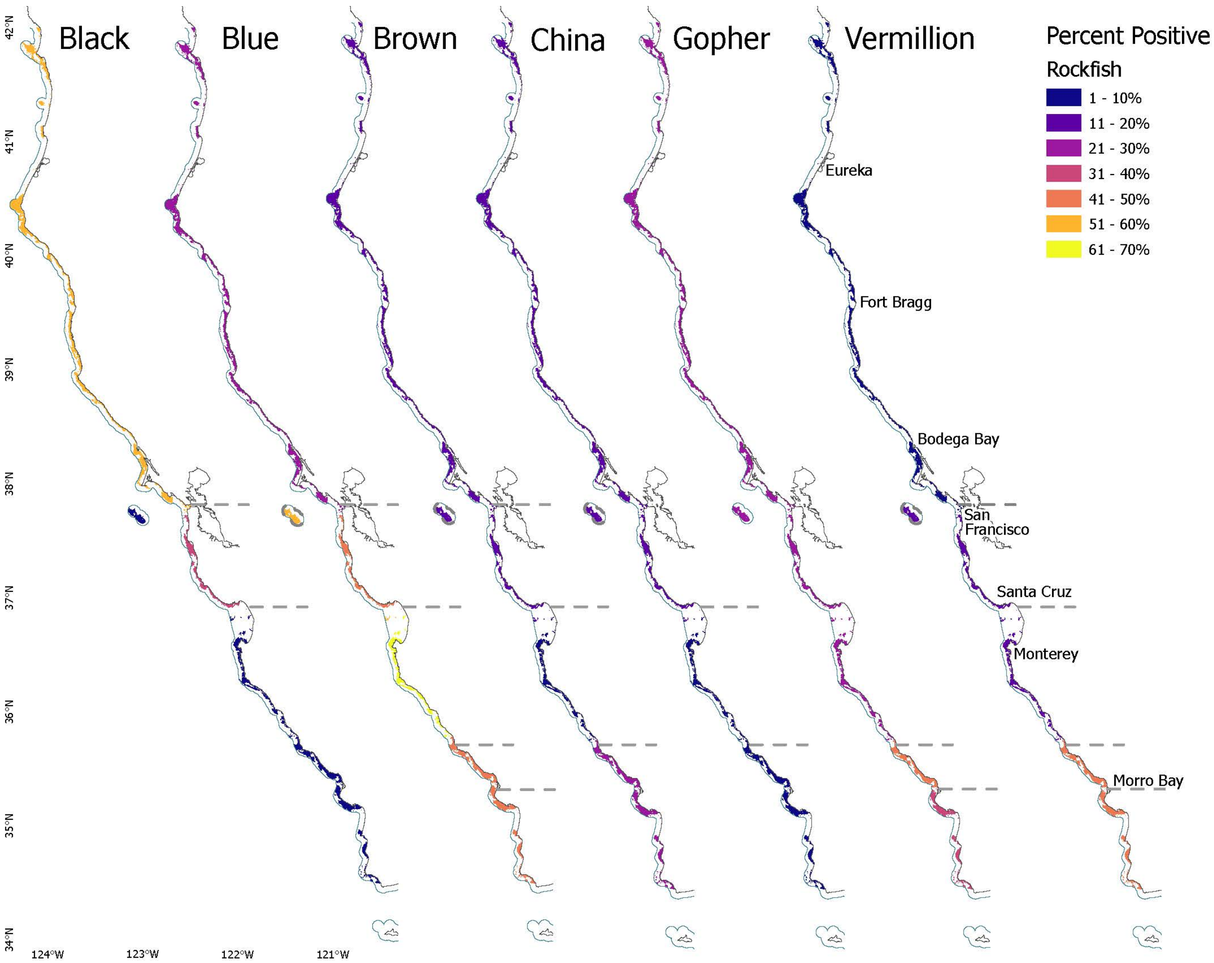


Figure 3: The percent of drifts that retained the target species, within grouped areas of rocky habitat over all years of the time series. The grey dashed lines represent the aggregated rocky habitat used to develop an index of abundance.

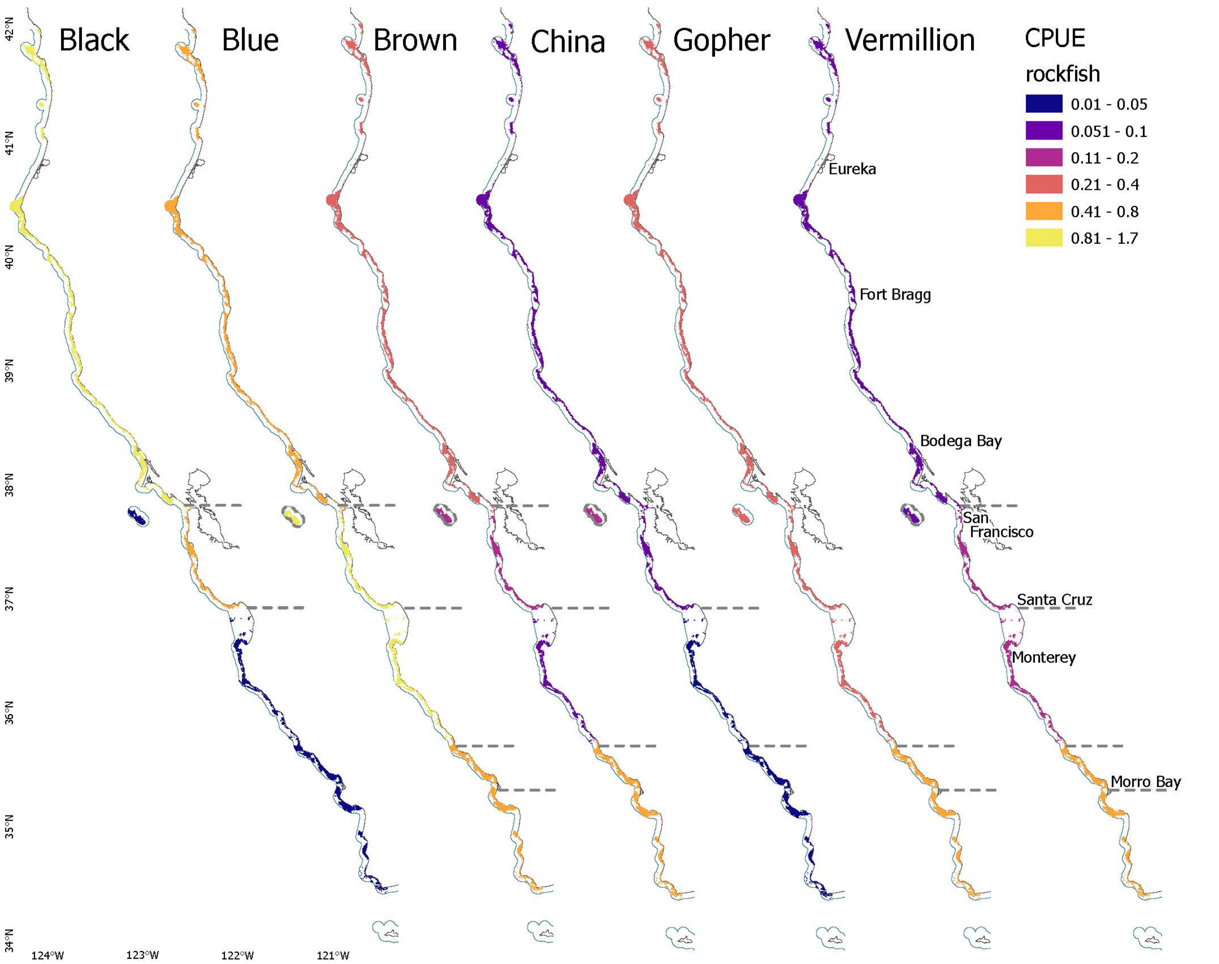
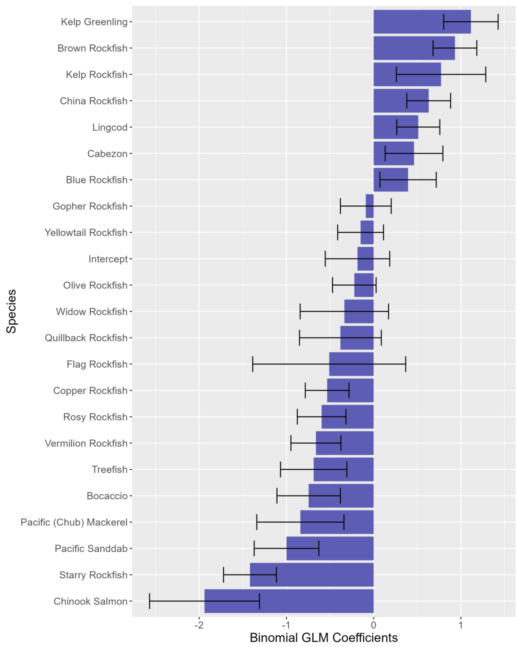
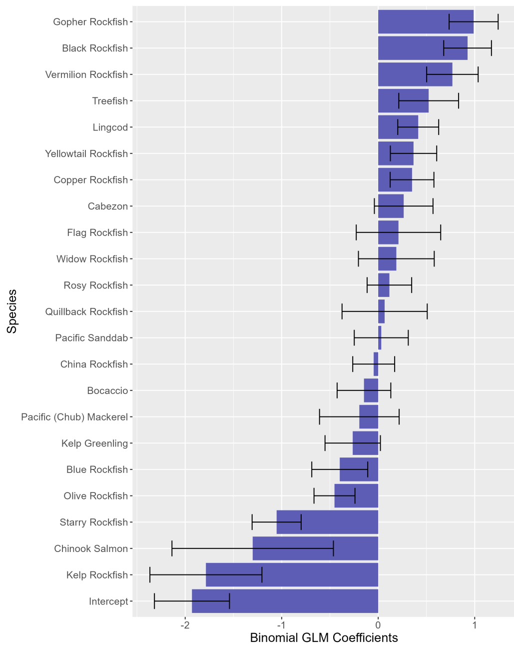


Figure 4: The average CPUE across all years of the time series for each of the six species. The grey dashed lines represent the aggregated rocky habitat used to develop an index of abundance.

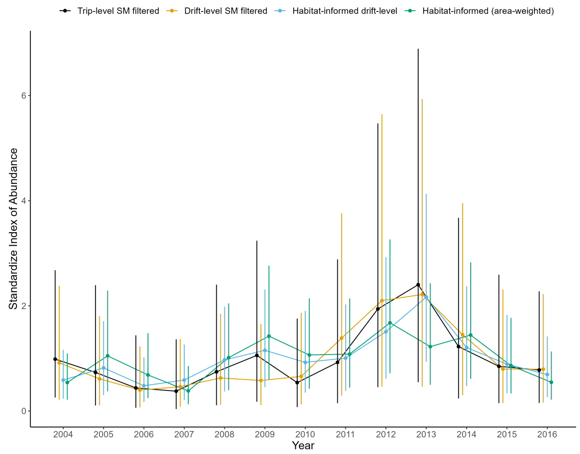


(a) Black rockfish trip-level

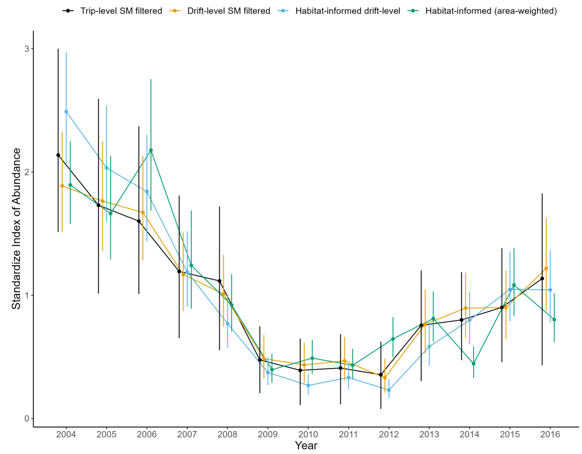


(b) Brown rockfish trip-level

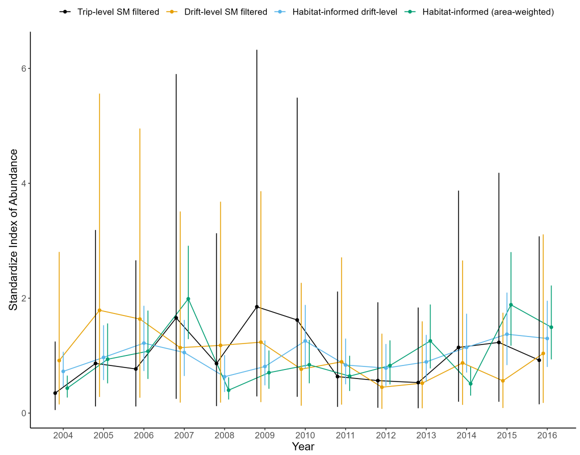
Figure 5: Examples of the species coefficients and 95% confidence intervals for the Stephens-MacCall filtering for black rockfish (a) and brown rockfish (b) in the trip-level data. A positive coefficient indicates a species is associated with the target species and a negative coefficient indicates the species is not associated with the target species.



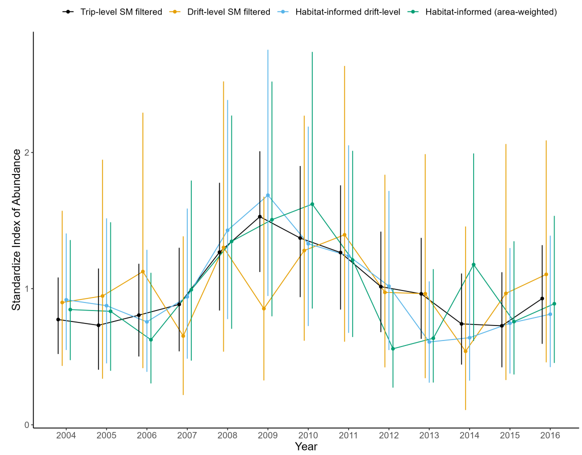
(a) Black rockfish



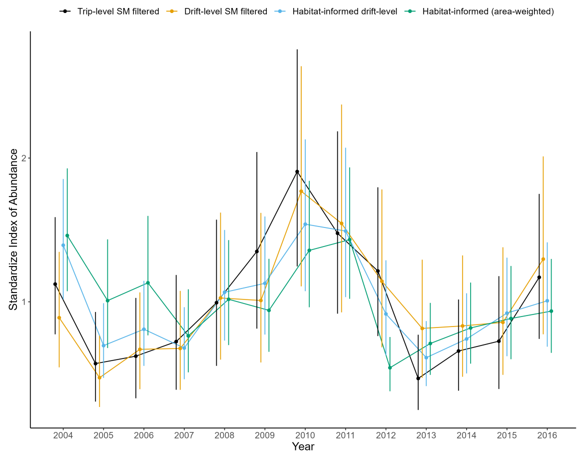
(b) Blue rockfish



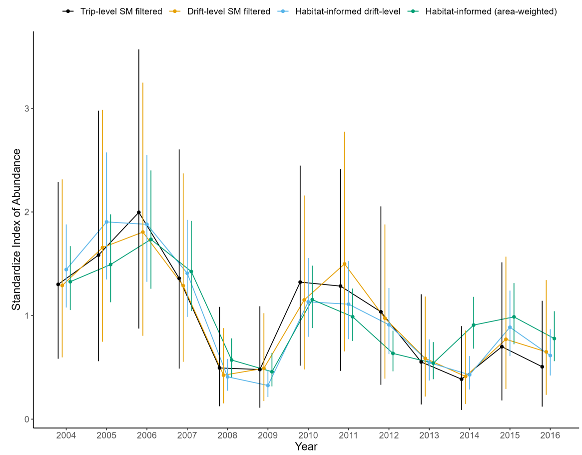
(c) Brown rockfish



(d) China rockfish



(e) Gopher rockfish



(f) Vermilion rockfish

Figure 6: Indices of abundance and 95% confidence intervals for the different filtering strategies, each scaled to its mean, for the six species.

References

Anderson, T.J., Yoklavich, M.M., 2007. Multiscale habitat associations of deepwater demersal fishes off central california. Fishery Bulletin 105.

Bacheler, N., Bulletin, K.S.F., undefined 2015, 2015. Estimating relative abundance and species richness from video surveys of reef fishes. sedarweb.org URL: http://sedarweb.org/docs/wsupp/S45_RD01_Bacheler%20and%20Shertzer%202014%20Species%20Abundance%20and%20Richness%20from%20Trap%20Surveys.pdf, doi:10.7755/FB.113.1.2.

Bay, S.M.L.C.S.U.M., 2014. California seafloor mapping project - undersea imagery archive 2007-2014. 137 blocks.

Campbell, R.A., 2015a. Constructing stock abundance indices from catch and effort data: Some nuts and bolts. Fisheries Research 161. doi:10.1016/j.fishres.2014.07.004.

Campbell, R.A., 2015b. Constructing stock abundance indices from catch and effort data: Some nuts and bolts. Fisheries Research 161, 109–130. URL: http://www.sciencedirect.com/science/article/pii/S0165783614002185, doi:https://doi.org/10.1016/j.fishres.2014.07.004.

Cope, J., Dick, E., MacCall, A., Monk, M., Soper, B., Wetzel, C., 2015. Data-moderate stock assessments for brown, china, copper, sharpchin, stripetail, and yellowtail rockfishes and english and rex soles in 2013. Pacific Fishery Management Council, Portland, OR .

Dettloff, K., 2021. Improvements to the stephens-maccall approach for calculating cpue from multispecies fisheries logbook data. Fisheries Research 242. doi:10.1016/j.fishres.2021.106038.

Ducharme-Barth, N.D., Shertzer, K.W., Ahrens, R.N.M., 2018. Indices of abundance in the gulf of mexico reef fish complex: a comparative approach using spatial data from vessel monitoring systems. Fisheries Research 198, 1–13. URL: http://www.sciencedirect.com/science/article/pii/S0165783617302965, doi:https://doi.org/10.1016/j.fishres.2017.10.020.

Golden, N., 2013. California state waters map series data catalog: U.s. geological survey data series 781.

Haggarty, D.R., King, J.R., 2006. Cpue as an index of relative abundance for nearshore reef fishes. Fisheries Research 81, 89–93. URL: https://www.

sciencedirect.com/science/article/pii/S0165783606002189?casa_token=GpbJVim4fAMAAAAA:uCaxpPua3ksx9pmTLM3hTa6sCjikRVQz0FVipwQaQ8HlRBX6qnaw00xvQc-DadDGXyJPpens7A, doi:10.1016/j.fishres.2006.05.015.

Harley, S.J., Myers, R.A., Dunn, A., 2001. Is catch-per-unit-effort proportional to abundance? Canadian Journal of Fisheries and Aquatic Sciences 58, 1760–1772. doi:10.1139/f01-112.

Hilborn, R., Walters, C.J., 1992. Quantitative fisheries stock assessment: choice, dynamics, and uncertainty. Springer Science and Business Media.

Laidig, T.E., Watters, D.L., Yoklavich, M.M., 2009. Demersal fish and habitat associations from visual surveys on the central california shelf. Estuarine, Coastal and Shelf Science 83. doi:10.1016/j.ecss.2009.05.008.

Larson, R.J., 1980. Territorial behavior of the black and yellow rockfish and gopher rockfish ({Scorpaenidae, Sebastes}). Marine Biology 58, 111–122. doi:10.1007/BF00396122.

Lo, N.C., Jacobson, L.D., Squire, J.L., 1992. Indices of relative abundance from fish spotter data based on delta-lognormal models. Canadian Journal of Fisheries and Aquatic Sciences 49, 2515–2526.

Love, M., Yoklavich, M.M., Thorsteinson, L., 2002. The rockfishes of the northeast Pacific. University of California Press.

Maunder, M.N., Punt, A.E., 2004. Standardizing catch and effort data: A review of recent approaches. Fisheries Research 70, 141–159. doi:10.1016/j.fishres.2004.08.002.

Miller, R.R., Field, J.C., Santora, J.A., Schroeder, I.D., Huff, D.D., Key, M., Pearson, D.E., MacCall, A.D., 2014. A spatially distinct history of the development of california groundfish fisheries. PLoS ONE 9. doi:10.1371/journal.pone.0099758.

Monk, M., Dick, E., Field, J., Saas, E., Rogers, T., . The status of vermilion rockfish (*sebastes miniatus*) and sunset rockfish (*sebastes crocotulus*) in u.s. waters off the coast of california north of point conception in 2021. Pacific Fisheries Management Council, Portland, Oregon. .

Monk, M.H., Dick, E.J., Pearson, D., 2014. Documentation of a relational database for the california recreational fisheries survey onboard observer sampling program, 1999-2011. NOAA-TM-NMFS-SWFSC-

529 .

Siegfried, K.I., Williams, E.H., Shertzer, K.W., Coggins, L.G., 2016. Improving stock assessments through data prioritization. Canadian Journal of Fisheries and Aquatic Sciences = Journal canadien des sciences halieutiques et aquatiques 73, 1703–1711. URL: https://cdnsciencepub.com/doi/full/10.1139/cjfas-2015-0398, doi:https://doi.org/10.1139/cjfas-2015-0398.

Stefánsson, G., 1996. Analysis of groundfish survey abundance data: Combining the glm and delta approaches. ICES Journal of Marine Science 53, 577–588. doi:10.1006/jmsc.1996.0079.

Stephens, A., MacCall, A., 2004. A multispecies approach to subsetting logbook data for purposes of estimating cpue. Fisheries Research 70, 299–310.

Valentine, J.W., 1966. Numerical analysis of marine molluscan ranges on the extratropical northeastern pacific shelf. Limnology and Oceanography 11, 198–211. doi:https://doi.org/10.4319/lo.1966.11.2.0198.