Fisheries Research 106 (2010) 298–309

Contents lists available at ScienceDirect

Fisheries Research

journal homepage: www.elsevier.com/locate/fishres

Analysis of fishery-independent hook and line-based data for use in the stock assessment of bocaccio rockfish (*Sebastes paucispinis*)

John H. Harms*, John R. Wallace, Ian J. Stewart

NOAA Fisheries, Northwest Fisheries Science Center, 2725 Montlake Blvd East, Seattle, WA 98112, USA

ARTICLE INFO

Article history:
Received 7 January 2010
Received in revised form 25 August 2010
Accepted 26 August 2010

Keywords:
Standardized catch per unit effort
Relative index
Hook and line research survey
GLM
Sebastes paucispinis

ABSTRACT

Fishery-independent surveys are an important source of information for stock assessment and management worldwide. Research surveys often use trawl gear to capture commercially valuable species and calculate indices of relative abundance or density. However, many species of interest do not occur in direct contact with the bottom, or occur in areas where high-relief habitat precludes trawl operation. This paper introduces a standardized hook and line survey for rockfish conducted by NOAA Fisheries' Northwest Fisheries Science Center in the Southern California Bight. The survey uses fishing gear similar to that used in many recreational fisheries to sample approximately 120 locations covering a wide range of depths and habitats. To provide an example of how these data can be analyzed for direct inclusion in stock assessments, we standardize catch rates of bocaccio rockfish from 2004–2008 using a Bayesian Generalized Linear Model to account for site, fishing time, survey vessel, angler, and other statistically significant effects. Results indicate that the bocaccio stock vulnerable to this survey in the Southern California Bight has shown a relatively flat trend over recent years. Length frequency distributions indicate the presence of several strong cohorts that should be detectable in future stock assessments of bocaccio for use in U.S. West Coast groundfish management. This survey is the only available tuning index for the adult portion of the bocaccio population in recent years as historically used recreational catch per unit effort indices have been compromised due to changes in bag limits and other management restrictions.

Published by Elsevier B.V.

1. Introduction

Fishery-independent surveys are crucial to successful assessment and management of fish stocks around the world. Trawl gear is efficient, readily standardized, and can provide information on a broad array of fishes. However, many species occur in areas that are not easily accessed with bottom trawls. Alternatives include acoustic surveys and visual observations collected from manned submersibles, remotely operated vehicles (ROVs), and autonomous underwater vehicles (AUVs). These methods can be an effective means of monitoring some species (e.g., Ressler et al., 2007; Yoklavich et al., 2007). However, acoustic surveys may not be feasible for monitoring demersal species close to the seafloor or where different species frequently occur in mixed schools. Visual surveys do not provide detailed biological data (e.g., sex, age, and genetic information which can aid in positive species identification and facilitate research into stock structure), and fish behavior may bias abundance estimates in unknown ways (Stoner et al., 2008).

On the west coast of the U.S., fishery-independent bottom trawl surveys have been conducted by the National Oceanic and Atmospheric Administration (NOAA) Fisheries since 1977 on the continental shelf areas between Canada and Mexico and since 1984 on the continental slope (Keller et al., 2006, 2007; Turk et al., 2001). However, of the approximately 89 species managed under the Pacific Fishery Management Council's (PFMC) Groundfish Fishery Management Plan (FMP), over 50% are encountered infrequently enough to make indices of abundance extremely imprecise. Many of these species including dozens of species of rockfish (genus: *Sebastes*) are common, but associated with rocky, complex habitats that are often poorly sampled using bottom trawls (Jagielo et al., 2003). Insufficient fishery-independent data for reliable monitoring of structure-associated species is a concern in many other fisheries and is not confined to the U.S. West Coast (e.g., SEDAR, 2007).

Stock assessments for these species have generally relied on indices of abundance from commercial or recreational fishery catch per unit effort (CPUE; e.g., bocaccio rockfish [*Sebastes paucispinis*], Field et al., 2009). However, problems associated with reconciling different bag limits, depth restrictions, and non-standardized methods and areas of fishing have led to the truncation of fishery CPUE time-series for nearly all managed species on the west coast of the U.S. in recent years. Fishery-independent data such as larval impingement indices and opportunistic sampling such as bottom trawl surveys conducted by the Los Angeles County and Orange

* Corresponding author. Tel.: +1 206 860 3414.
E-mail address: John.Harms@noaa.gov (J.H. Harms).

0165-7836/$ – see front matter. Published by Elsevier B.V.
doi:10.1016/j.fishres.2010.08.010

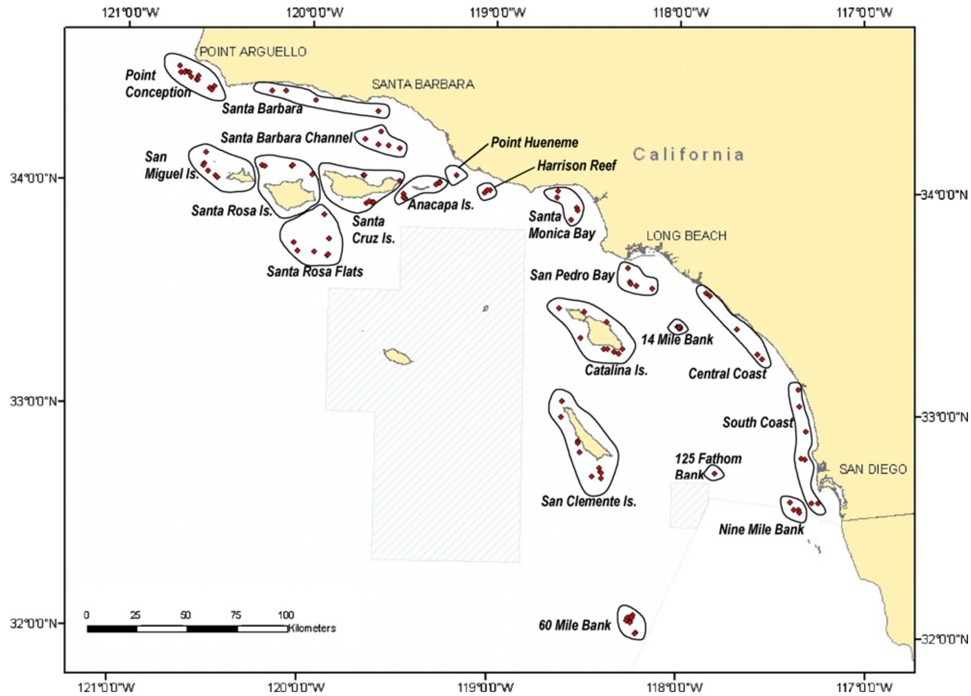

Fig. 1. Map of the Southern California Bight showing individual sampling sites and the 20 sampling areas. The boundary between the U.S. and Mexican Exclusive Economic Zones is indicated by the change in shading in the lower right portion of the map. The hatched areas indicate the two Cowcod Conservation Areas.

County sanitation districts are extremely geographically restricted and may do a relatively poor job of representing rockfish stocks over their larger range. Independent reviews of recent data collection procedures highlight the need for fishery-independent indices of abundance and have recommended expanded survey coverage in untrawlable habitats and the use of alternative sampling methods including visual observations, acoustic surveys, and hook and line-based gear (Pacific Fishery Management Council, 2006; Seger, 2000; US General Accounting Office, 2004).

In 2002, the Northwest Fisheries Science Center (NWFSC), Southwest Fisheries Science Center (SWFSC), Pacific States Marine Fisheries Commission (PSMFC), and local sportfishing industry collaborated to develop a hook and line survey to address the paucity of data for groundfish associated with untrawlable habitats in general and for Southern California in particular. Pilot studies were conducted in 2003, and a time-series of CPUE data and biological information for a variety of shelf rockfish species was begun in 2004. The survey's primary objective is to provide annual monitoring coverage of structure-associated shelf rockfish populations across the entire Southern California Bight (SCB) from Point Arguello (34°30″N) to the southern boundary of the U.S. Exclusive Economic Zone (EEZ) (Fig. 1).

In this paper we introduce the results from the first 5 years of the hook and line survey with two main objectives: (1) develop a general framework for standardizing hook and line survey observations into species-specific annual indices of abundance; and (2) apply this method to bocaccio rockfish (*S. paucispinis*) in Southern California. Bocaccio is of particular interest because it is one of seven overfished species that is currently limiting commercial and recreational fishing opportunities on the U.S. West Coast (MacCall, 2007a). Bocaccio are frequently encountered and occur over the entire geographic range sampled during the hook and line survey. Recent bocaccio stock assessments have lacked ongoing fishery-independent abundance indices for post-juvenile stage individuals.

2. Methods

2.1. Survey design

The NWFSC hook and line survey samples shelf rockfish occurring over hard substrate within the SCB across a relatively narrow band of depths (37–229 m) encompassing the common depth range for bocaccio (Love et al., 2002). This area of interest is markedly smaller than the total geographic area of the SCB. Despite this, comprehensive habitat maps with the level of resolution needed to accurately identify the spatial extent of all rockfish habitat are currently unavailable. Therefore, this survey developed a database of potential sampling sites through discussions with sport and commercial fishermen throughout the region, previous groundfish monitoring programs operated by California Department of Fish and Game (CDFG), and sites sampled opportunistically during the 2003 pilot study and other cruises. To ensure sampling coverage throughout the entire SCB in each year, the region was subdivided into 20 sampling areas (Fig. 1), and representative sites were selected from within each of these areas in proportion to the amount of potential rockfish habitat the area was judged to contain. Selected survey sites were distributed among areas except that in some areas all available sites were included. This subset of sampling locations represents a fixed-station design because not all possible (but unknown) sites have an equal probability of being sampled. The Cowcod Conservation Areas (CCAs) are currently excluded from the sampling design. Sites were selected to include the range of habitats and fishing histories (popular charter boat fishing locations as well as relatively obscure rockpiles) likely to be important throughout the region (Harms et al., 2008). Oil platforms and other artificial habitat features were not included in the sampling database.

Surveys were conducted annually between 2004 and 2008 with two commercial passenger fishing vessels (CPFVs) chartered for 10–12 days each to complete all sampling activities. Since 2005,

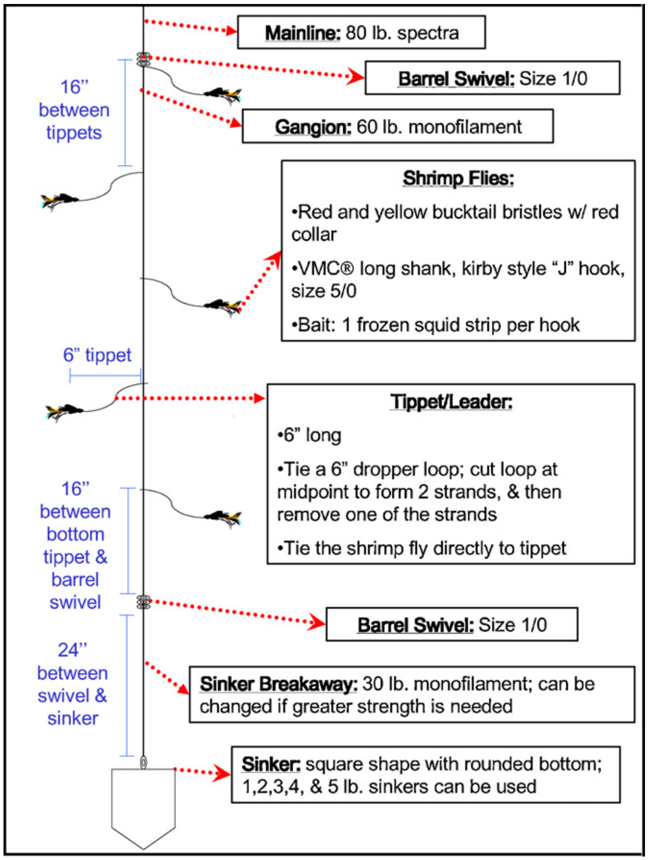

Fig. 2. Diagram of gangion used during hook and line survey.

each survey has begun during late September and ended in early October (Table 1). In 2004, the survey was conducted in November due to vessel availability. During the first 5 survey years, 121 unique sites were sampled, with 120 of those visited in the most recent year's survey. Sites are sampled in a single visit by a single vessel and are not sampled again until the next survey year. Early survey years included refinement of the sampling frame to remove sites where suitable rockfish habitat is not present. These refinements were based on camera and acoustic observation, not observed catch rates. Although all sites within the sampling frame contain potential rockfish habitat, not all sites would be characterized as optimal bocaccio habitat as at least 20% of all sites sampled in every year have yielded no bocaccio catch (Table 1).

2.2. Sampling methods

Sampling was conducted using hook and line gear deployed by rod and reel. Upon arrival at a vessel's designated sampling site, three deckhands each made five concurrent drops of a vertically arranged, five-hook sampling rig (or gangion), providing for a maximum possible catch of 75 fish per site. All five drops are generally completed in approximately 1 h. The sampling gangion was developed with input from the sportfishing industry to be appropriate for shelf rockfish in the region and similar to the recreational gear historically used in the area (Fig. 2). Electronic stopwatches were used to measure the amount of time the gangion spends on the seafloor available to demersal rockfish as well as when the first bites are observed on each drop. The maximum allowable soak time (hereafter, "fishing time") per drop was 5 min; however deckhands had discretion to reel in the line at any point prior to this to avoid hooked fish slipping free from the gangion. All time measurements were recorded as they were collected. In rare instances,

Table 1
Summary of hook and line sampling activity and bocaccio catches, 2004–2008.

	2004	2005	2006	2007	2008	Sum
Survey dates	10–21 November	27 September–8 October	26 September–8 October	25 September–7 October	25 September–8 October	n/a
Number of sites	75	90	92	99	120	476
Number of bocaccio	791	669	745	649	667	3,521
Weight (kg) of bocaccio	1121.4	1097.1	1053.6	881.1	955.6	5108.8
Bocaccio catch (% of total bocaccio optimum yield[a])	0.56%	0.36%	0.34%	0.40%	0.44%	n/a
Proportion sites with bocaccio captures	78.7%	76.7%	79.3%	79.8%	75.8%	78.1% (mean)
% hooks capturing fish	43.4%	37.0%	35.1%	35.4%	35.6%	37.3% (mean)
% hooks capturing bocaccio	14.4%	10.2%	11.2%	8.9%	7.5%	10.4% (mean)

[a] Amount of fish that provides the greatest overall benefit to the nation and provides for rebuilding of overfished species. Optimum yields are set biennially by the PFMC.

Table 2
Description of categorical and continuous explanatory variables evaluated during model selection.

Group	Variable	Description
Sampling	Angler position	Categorical: 1 = bow; 2 = midship; 3 = stern
	Drop number	Categorical: 1 = first drop; through 5 = last drop
	Hook position	Categorical: 1 = bottom hook; through 5 = top hook
	Vessel	Categorical: vessel 1; vessel 2
	Fishing time	Continuous: elapsed time between gangion reaching bottom and beginning retrieval (seconds)
Location	Site	Categorical: 74–120 unique observations depending on year
	Latitude[a]	Continuous: latitude at each drop (decimal degrees)
	Depth[a]	Continuous: bottom depth for each drop (meters)
	Bathymetric relief[a]	Continuous: difference between deepest and shallowest drops ever conducted at the same site (meters)
	Distance to major port[a]	Continuous: distance from site to the nearest major fishing port (i.e., Santa Barbara, Ventura, Oxnard, Marina Del Rey; San Pedro, Long Beach, Newport Beach, Dana Point, Oceanside, Mission Beach, San Diego) (kilometers)
	Distance to centroid	Continuous: distance from each drop to the calculated centroid of all drops ever conducted at a site (meters)
Ocean/weather	Swell height	Continuous: visually calculated for each site (meters)
	Swell direction	Continuous: determined visually with compass for each site (degrees)
	Wave height	Continuous: visually calculated for each site (meters)
	Drift speed	Continuous: calculated using global positioning system (GPS) (nautical miles per hour)
	Drift direction	Continuous: determined with GPS for each site (degrees)
	Wind speed	Continuous: estimated by manual observation (nautical miles per hour)
	Wind direction	Continuous: determined with compass for each site (degrees)
	Sea surface temperature	Continuous: determined using vessel-mounted thermometer (degrees Centigrade)
Temporal	Year	Categorical: year effect as it relates to abundance
	Time of day	Continuous: 24 h clock
	Tide phase	Categorical: as reported by closest tide station (ebb; flood; steady)
	Tide height	Continuous: as reported by closest tide station (meters)
	Tide type	Categorical: spring; neap; neither
	Moon phase	Continuous: 29.5-day lunar calendar

[a] These covariates are defined at the site level and therefore not included in the final model which contains site.

the maximum fishing time was exceeded due to gear problems or other logistical issues. Captured fish were identified to species, and length, weight, and sex were recorded. Both sagittal otoliths were extracted for laboratory ageing, and a fin clip was taken for future genetic analyses.

2.3. Standardization via a Generalized Linear Model

Generalized Linear Models (GLMs) are a standard tool for converting raw catch and catch-rate data into standardized indices of abundance for use in stock assessments (Maunder and Punt, 2004). Although we explored the use of aggregated observations of catch and catch rate (e.g., fish per site, fish per hook-hour, etc.), we model the data at their most basic level which is directly in the form they are collected: the individual hook. Any hook deployed in the survey design either captured the species of interest, or it did not, leading logically to a binomial error structure. Each hook is a unique intersection of survey design-related variables (e.g., year, site, vessel, angler position, drop number, and hook position) and represents an independent observation of the bocaccio resource within a GLM framework. Effort (in this case, duration of fishing during each drop) can be used to standardize raw catch into a catch rate, but this requires careful examination of the proportionality prior to standardization (Xiao, 2004). A simpler approach is to model catch, instead of catch rate, and allow one or more metrics of effort to be included in the GLM as offsets or explanatory variables. We take this approach, allowing the total fishing time of each gangion to define the effort expended; this allows the proportionality between hooks and time to be estimated in the GLM.

Adding many explanatory variables to this type of model will generally reduce bias in the estimated year effects, at the cost of increased variance (Maunder and Punt, 2004). We therefore pre-selected a subset of potential covariates (Table 2) based on experience on the survey and inspection of collinearity (although collinearity itself may be less important in the context of developing a standardized index than model fitting where the coefficients themselves are the results of interest). The categorical variables related to survey design (e.g., vessel, site, angler position, drop number, hook position) were modeled as fixed, rather than random effects. This is a less restrictive approach that makes no distributional assumption about the relationships among estimated coefficients. Therefore, the results of these analyses represent the estimated trend for a fixed set of locations sampled as described and do not include bocaccio abundance outside of the survey's sampling frame. Modeling the data with a standard binomial GLM at the drop, site, or other higher levels of aggregation incorrectly treats hooks as repeated observations and implicitly assumes each hook has an equal probability of capturing a fish, regardless of its position on the gangion, the position on the vessel from which it is deployed, or (modeling at the site level) whether it is deployed early or late during the sampling visit. This could be especially problematic in situations where gear or logistical problems are encountered (e.g., broken hooks and snagged lines) or where sampling must be completed before all five drops have been deployed (e.g., sudden weather changes, military operation in the area, etc.).

Although it may be appealing to use the presence and capture of co-occurring species as covariates to account for saturation and other effects, this is inappropriate in this instance because similar patterns of recruitment or mortality may cause these covariates to explain a portion of the inter-annual trend (Maunder and Punt, 2004). For this reason, we do not include other species' catch rates in the analysis; however we do evaluate observed levels of gear saturation for the potential for interaction among catch rates. Exploration of continuous variables via generalized additive models (GAM) was performed to evaluate whether linear, simple polynomial or more complex functional forms were warranted. Although the use of polynomial transformations is often discouraged during standardization of catch and effort data (Maunder and Punt, 2004), it may be necessary where covariates have a distinctly non-linear relationship with catch rate.

From the *a priori* subset of explanatory variables, we then employed DIC (Deviance Information Criterion; Spiegelhalter et al., 2002) based on the posterior distributions of the relative likelihood for each model including main effects, 2nd degree polynomials of the main effects, and simple interaction terms in sequence in the linear predictor. DIC is a Bayesian goodness-of-fit metric interpreted in a similar fashion to likelihood-based criteria such as AIC (Akaike's Information Criteria; Burnham and Anderson, 2002) or BIC (Bayesian Information Criterion; Schwarz, 1978).

Link functions appropriate to the binomial error model, including the canonical 'logit', the 'probit', and the 'complementary log–log', were evaluated for preliminary models including all main effects but no interactions. The best model, given each of the different candidate link functions, was found via a stepwise procedure using both forward and backward selection.

The GLM with a logit link function can therefore be represented as:

$$\mu_{i,j,k,l,m} = \text{logit}^{-1}(\theta_{i,j,k,l,m,n})$$

for i years, j sites, k drops, l vessels, m anglers, and n hooks, and where θ is the linear predictor. The probability of capture for a bocaccio on each hook is:

$$\text{bocaccio} \sim \text{Bernoulli}(\mu_{i,j,k,l,m,n})$$

All GLMs were fit using the R statistical package (R Project, 2009), and the Bayesian posteriors were sampled with the function 'MCMClogit' contained within the 'MCMCpack' R package (MCMCpack, 2009). All survey sites are included in the analysis (including those that register zero bocaccio catch), however the data used to fit the GLM includes only those sites that were sampled in at least 2 years and for which at least one bocaccio was captured. This approach implicitly assumes that the estimated probability of capture is proportional to the abundance of fish. Although this cannot possibly be true for all levels of abundance given the fixed and finite amount of survey effort deployed, we evaluate the assumption based on the observed levels of gear saturation (see Section 4).

Comparison of observed and expected catches, and inspection of residual patterns at several levels of aggregation were performed in order to evaluate the GLM approach for overdispersion, systematic patterns in lack-of-fit, and ability to adequately explain the observed data. Convergence of the Bayesian integration to a stationary distribution for all model parameters and the standardized index of abundance was assessed using several well established diagnostic tools widely available through the "coda" package in R. For each quantity, cumulative quantiles of the posterior distribution, maximum autocorrelation at lags to 20, the Geweke statistic and the Heidelberger–Welch statistic were calculated.

2.4. Calculation of index and variance for use in stock assessments

It is common practice to back-transform the linear predictor using the inverse link function in order to capture the appropriate variance components (including that of all year effects) in the final index of abundance. Further, if a link other than the identity link is used, the back-transformed values occur in the space of the original observations (not in the model space, as do coefficients themselves). To perform this back-transformation, levels of the other explanatory variables (i.e., all those except the variable of interest) must be selected at standardized values. In this case, we used the median of the two continuous variables and the first level of each factor (set to zero in the GLM) except for vessel; back-transformed indices were averaged over both vessels for easier interpretation. The year coefficients are individually retained during back-transformation to create the time-series. For the binomial model with logit link function, the standardized back-transformed index can be interpreted as the probability of capturing a bocaccio on a single hook in each year given a fixed value for all other explanatory variables. Alternate back-transformations were performed with the explanatory variables set to different standardized values to explore the possibility that the initial levels selected influenced the variability of the indices (Maunder and Punt, 2004).

Analytic calculation of the variance of the standardized index is not straightforward. However, Bayesian integration via Markov chain Monte Carlo (MCMC) easily propagates the variance of all model coefficients into the back-transformed index (the quantity of interest). We selected noninformative conjugate priors for model parameters and performed sensitivity tests to ensure the results were not heavily influenced by the choice of these priors. Appropriate convergence diagnostics were applied to ensure the MCMC chain had reached stationarity.

2.5. Power analysis

We used a power analysis to explore the ability of the method developed in this paper to identify meaningful changes in the bocaccio population given the survey design, the observed sample sizes through 2008, and the GLM-based standardization model. To do this, we simulated data sets for one future year of sampling, 2009, with the percentage of bocaccio captured randomly increased or decreased (relative to 2008) over a range from 5% to 50%. Each simulated data set was then re-analyzed, along with the existing data, using the same GLM. The posterior distributions for the 2008 index and the simulated estimate for 2009 were compared using an alpha value of 0.10 for Type 1 error. This approach specifically evaluates the question: What is the power of detecting a simulated change in abundance between 2008 and 2009 while incurring no more than a 10% chance of concluding that there is a change in abundance, when there is not? This method was computationally intensive, as a full Bayesian integration was completed for each simulated data set. However, a minimum of 30 simulations were conducted for each level of change in percent bocaccio catch (±5%, 10%, 20%, 30%, 40%, and 50%) in order to ensure that the central tendency of the power relationship was stationary.

3. Results

In 2004, 75 sites were successfully sampled (Table 1). The number of sites sampled has increased each year, and by 2008, data were collected at 120 sites. Of the 121 unique sites that have been visited in 2004–2008, 44 have been sampled in each survey year, 46 in 4 years, 11 in 3 years, and 18 in 2 years. Two sites have been sampled in only 1 year and are not included in the current analysis, but may be added in the future if they can be revisited.

Between 35.1% and 43.4% of all valid[1] hooks deployed in any survey year have yielded fish (Table 1). Bocaccio have been hooked on 7.52–14.4% of all valid hooks deployed in a survey year, and at least one bocaccio has been caught at over 75% of the sites in each year (Table 1). Bocaccio have accounted for 21.2–33.1% by frequency and 28.7–41.8% by weight of all fish caught on valid hooks during the first 5 years of the hook and line survey. From 2004 to 2008, raw catch rates ranged from 5.9 to 10.9 bocaccio caught per site. "Floater" fish (i.e., fish that are hooked at depth but float to the surface free of the gangion during retrieval) are rare (<0.3% of all fish hooked) and are randomly assigned to available hooks. The two measures of raw catch rates are variable, and have not shown a clear

[1] Valid hooks exclude damaged or missing hooks; hooks from gangions that become badly tangled or snagged on bottom; and any remaining hooks on gangions that return with three or more missing or damaged hooks. Valid hooks have accounted for at least 96.4% of all hooks deployed in each survey year.

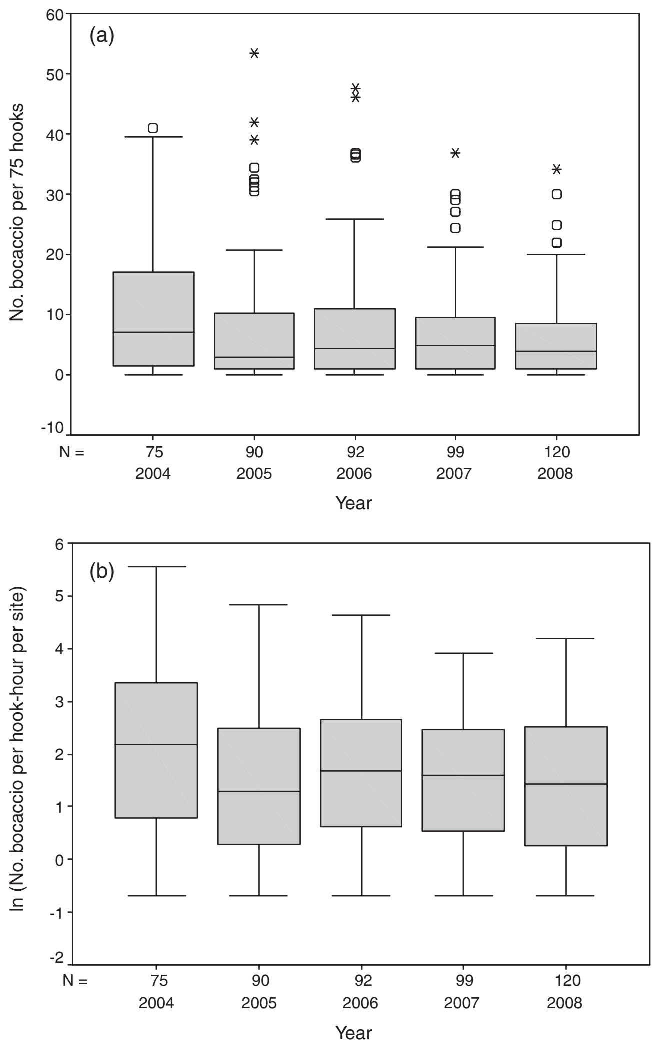

Fig. 3. Raw catch rates for two metrics of raw bocaccio catch rate, 2004–2008: number of bocaccio captured per 75 hooks (panel a), and log-transformed number of bocaccio captured per hook-hour per site (panel b).

increasing or decreasing trend in abundance since 2005 (Fig. 3a and b).

The final linear predictor selected for the GLM was:

$$\theta_{i,j,k,l,m,n} = \alpha + \text{Year}_i + \text{Site}_j + \text{poly(Fishing time, 2)} + \text{Vessel}_l + \text{Hook position}_n + \text{Angler position}_m + \text{Drop number}_k + \text{poly(Wave height, 2)} + \text{Drop position}_k \cdot \text{Angler position}_m$$

where a "·" represents an interaction term, and poly() identifies 2nd degree polynomials for continuous variables. Contribution to explained deviance and reduction in DIC for all selected variables is described in Table 3. With the possible exception of the interaction of vessel and tide type, all of the variables selected for the final model have a clear interpretation of their influence on bocaccio catch during the survey (Table 3 and Fig. 4). The negative correlation between catch and fishing time (Fig. 4) appears counterintuitive *vis-à-vis* the generally positive relationship observed in many longline surveys. These surveys largely depend on attracting additional fish to available hooks via an olfactory response to a gradually expanding bait plume (Sigler, 2000). In contrast, the hook and line survey directly targets aggregations of fish, so strikes are more likely to be a visual response to a gangion dropped in very close proximity. Further, as deckhands are granted discretion in retrieving their lines, they often base this decision upon how quickly they perceive bites on the gangion. In areas of high abundance and prompt fish strikes, deckhands tend to initiate retrieval very quickly to prevent hooked fish from slipping free; however, in lower abundance areas where strikes are less plentiful or absent, deckhands may wait the maximum allotted 5 min with the hope that a fish might be encountered before concluding the drop Additionally, the small, positive-sloping shoulder of the relationship between catch and fishing time is obscured by the second-degree polynomial's fit to the larger number of observations at longer fishing times.

The binary response and predictions were aggregated such that residuals could be evaluated at several scales. Due to the nature of fitting a binomial model of this size, aggregating at any single factor results in a perfect fit, hence showing the residuals by site and year appeared to be a reasonable compromise. The site-level residuals (Fig. 5) show little evidence of bias, and generally conform to the binomial error structure, where the variance increases as a function of the expectation and sample size. An estimate of the model dispersion was derived by applying a quasi-likelihood fitting; the value was 0.9733, further indicating that the binomial error structure was generally appropriate for this problem.

Convergence of the posterior distribution was achieved by running the MCMC algorithm for 20 million iterations, removing the first 5 million and then retaining every 15,000th iteration, for a total of 1000 samples. For all coefficients and index estimates, the maximum (absolute value) autocorrelation at lags 1–20 was no more than 11.5%, and inspection of the cumulative quantiles indicated that even the 5th and 95th were stable. The distribution of the Geweke statistic, testing for stationarity of the central tendency, was distributed approximately standard normal, and no quantity failed the Heidelberger and Welch statistic.

Table 3
Component table for the final bocaccio model, in order of reduction in ΔDIC, except for 'Year' which was kept as the first variable. The term 'poly() identifies 2nd degree polynomials.

	Degrees of freedom	Explained deviance	*P* (*Chi*)	ΔDIC
Grand mean (α)	1	21,786.8	<0.001	3396.3
Year	4	199.1	<0.001	3205.1
Site	106	2532.1	<0.001	884.8
Poly(Fishing time)	2	583.3	<0.001	403.5
Vessel	1	28.1	<0.001	246.2
Poly(Tide height)	2	60.8	<0.001	174.0
Hook position	4	58.2	<0.001	121.9
Poly(Wave height)	2	49.3	<0.001	77.1
Angler position	2	15.0	<0.001	45.0
Drop number	4	78.6	<0.001	22.3
Tide type and/or interaction (vessel, tide type)[a]	4	28.1	<0.001	0.0

[a] Due to the dropping out of collinear columns, the model including the main effect of tide type and the interaction of vessel and tide type is identical to the model with just the interaction of vessel and tide type.

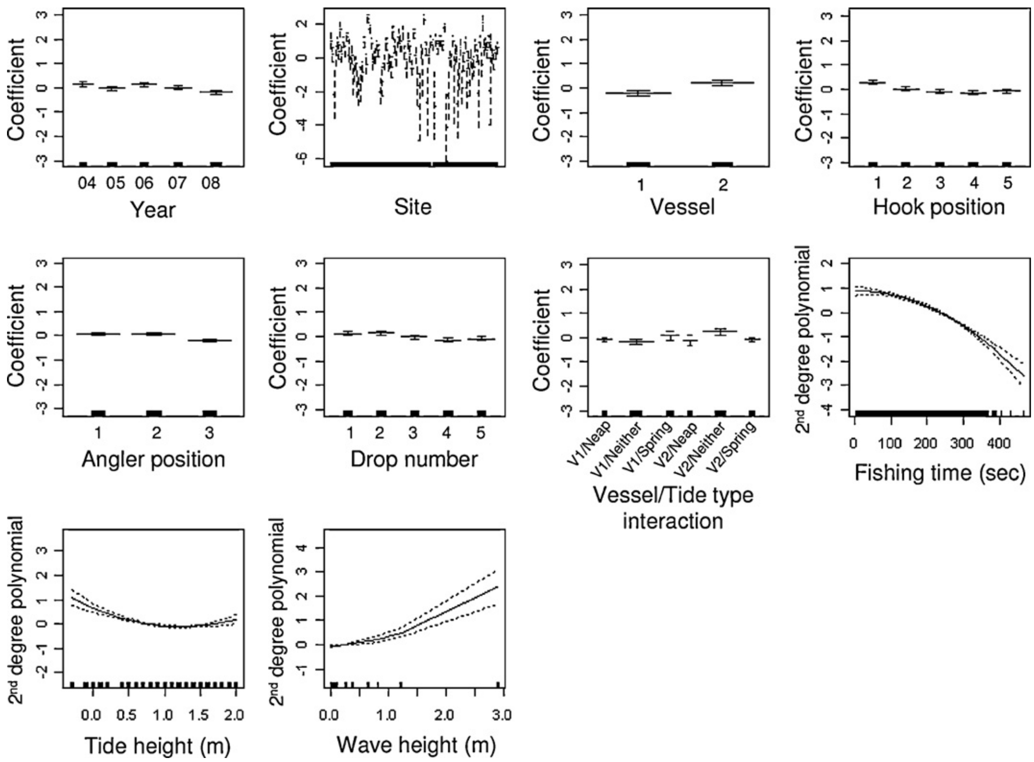

Fig. 4. Estimated model coefficients (±2SE) for the final selected bocaccio GLM.

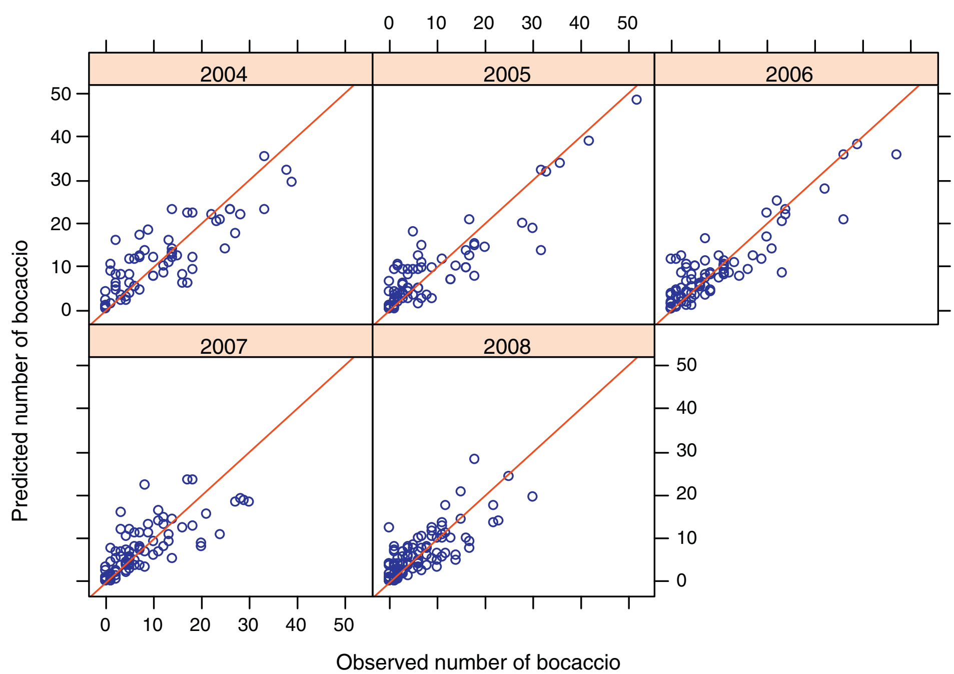

Fig. 5. Observed vs. predicted bocaccio catch by site and year. Lines indicate a 1:1 relationship.

Table 4
Standardized index values (back-transformed model predictions) for 2004–2008 and 95% credible intervals for the posterior median index (MCMC). Also presented for comparison are the year coefficients and standard errors directly from the GLM. Note that coefficients are on the scale of the linear predictors and are not directly comparable to the indices.

	2004	2005	2006	2007	2008
Posterior median index	0.1944	0.1644	0.1864	0.1696	0.1432
95% credible interval	0.131–0.280	0.106–0.245	0.122–0.274	0.112–0.249	0.092–0.214
Maximum likelihood index	0.1684	0.1422	0.1619	0.1463	0.1238
GLM year coefficient	0.00	−0.1999	−0.0468	−0.1672	−0.3600
Coefficient SE	0.00	0.0779	0.0704	0.0678	0.0708

Estimated year effect coefficients from the final model fit (Fig. 6a) and back-transformed predicted values (Fig. 6b) show a generally flat trend in abundance. To obtain the predicted values, we used the median of the two continuous variables (250 s for fishing time and 1 foot [0.30 m] for wave height) and the first level of each categorical variable (set to zero in the GLM). Standardizing the continuous variables to different values resulted in changes in the scale, but not the relative trend or uncertainty of the indices relative to the initial back-transformation. Removing marginally significant terms from the model had little effect on the final estimates, but appreciably reduced overall variance. The 95% credibility intervals shown in Table 4 are generally wide, reflecting the relatively short time-series and large number of relatively imprecisely estimated coefficients for this survey.

To directly illustrate the tradeoff between potential bias and precision, a simple model including only year as a covariate, and a second alternative model aggregating all catches to the site level were explored. The simple model (Fig. 6c) grossly underestimated the variance of each year's index estimate with CVs on the order of 3% (extremely small relative to other fishery-independent surveys; e.g., Helser et al., 2007). In general, such high confidence could lead to the identification of spurious trends if such a model were used when, in fact, there is a higher degree of uncertainty. The aggregated model provides an index and variance similar to that of the full model, indicating that, for bocaccio, potential sources of bias in this relatively short time-series are not severe (Fig. 6d). However, this may not be the case for future years or potential indices generated for co-occurring species.

The survey was found to have a reasonable degree of power to detect changes in relative abundance between 2008 and 1 year of simulated future data. For an increase or decrease in abundance of 30–40%, there was approximately a 60% probability of detection. However this relationship was asymmetric, with slightly greater power to detect decreases in relative abundance than an increase (Fig. 7).

4. Discussion

This analysis provides a novel statistical approach for modeling hook and line data to generate a fishery-independent index of abundance. The multiple factors of the survey's design (e.g., site, vessel, angler position, drop number, and hook position) yield a large and complex data set that does not intuitively lend itself to traditional modeling techniques and variance estimates. However, with the now widely used GLM methodology and in particular the option to use a binomial error structure, presence/absence data such as those modeled here can be analyzed in a logical framework.

The U.S. West Coast bocaccio resource was designated as overfished by NOAA Fisheries in 1999, and subsequent analyses have established a target year for rebuilding of 2026 (MacCall, 2007b). The hook and line survey was developed to track the rebuilding of the bocaccio population in Southern California and monitor the status of several other species of shelf rockfish. Although rod and reel gear is not commonly used to conduct research surveys, it provides an inexpensive, versatile means of effectively sampling demersal rockfish species associated with areas of structure and hard bottom. The high proportion positive site values for bocaccio coupled with the large range of observed catch per site data suggest the survey's fixed sites encompass appropriate habitat for the target species and that the sites represent a variety of levels of abundance. Length frequency distributions for bocaccio caught during the survey show clearly defined modes corresponding to known strong year classes and enable those cohorts to be tracked over time, suggesting generally constant selectivity for a wide range of sizes (Fig. 8). The survey's annual coverage throughout the SCB and its ability to pro-

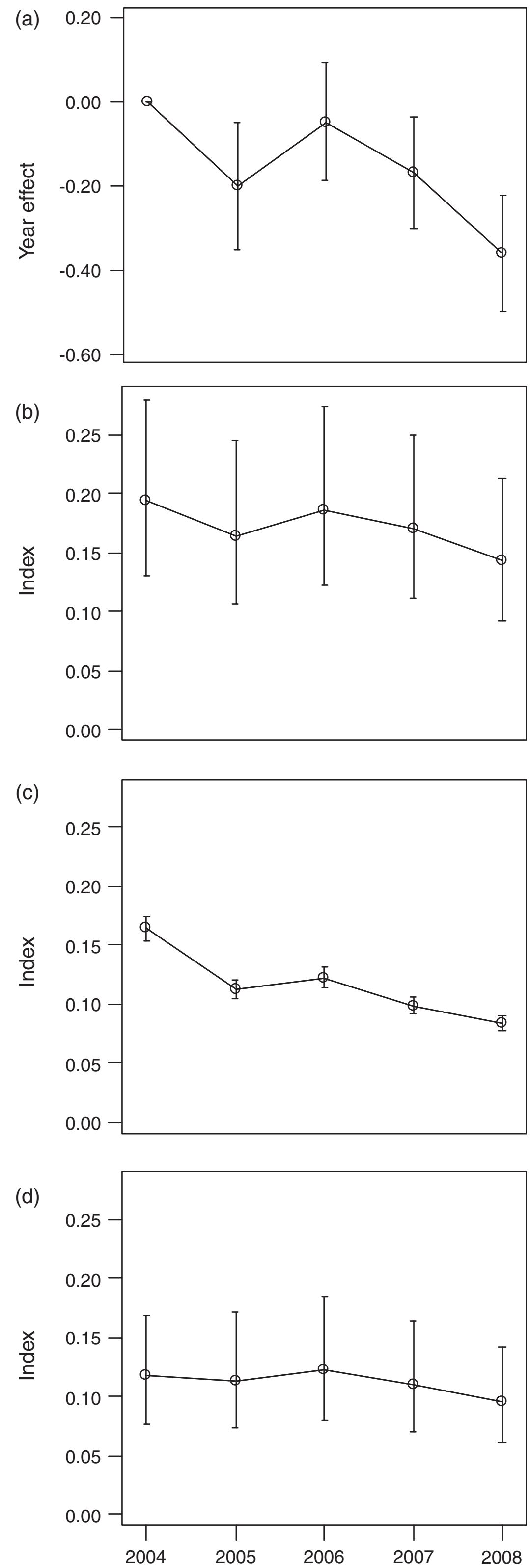

Fig. 6. Estimated values and 95% confidence intervals for the year effect coefficients (panel a); back-transformed posterior median index with 95% credible intervals (panel b); as well as alternate models including year as the only covariate (panel c); and fitting to data aggregated at the site level (panel d). Note the scale and space of year coefficients differs from the index values (linear predictor vs. response) and only the relative trend is comparable.

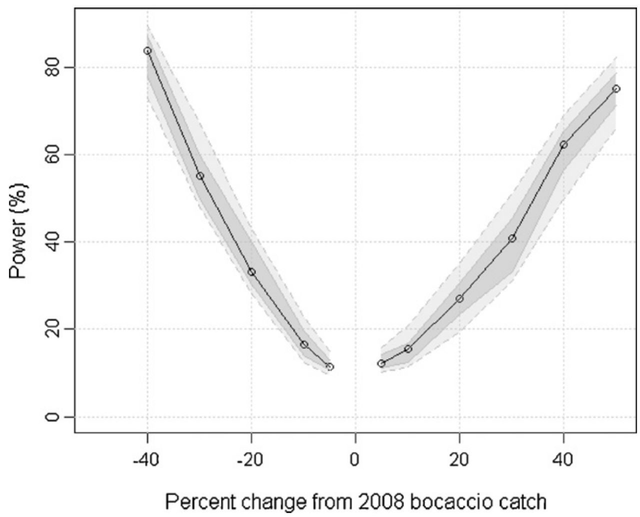

Fig. 7. Power analysis of the percent change in bocaccio catch for a Type 1 error alpha level of 10%. The connected circles are the median over all simulations; the dark grey area represents the range of the 25th and 75th percentiles and the light grey area the 10th and 90th percentiles.

vide sex and precise length observations along with otoliths for age and growth estimations are useful for reducing uncertainty in stock assessments.

All fishery-independent abundance indices currently available for use in bocaccio stock assessments are imprecise. The 2007 bocaccio stock assessment update (MacCall, 2007a) recommended that the triennial (see Wilkins et al., 1998) and NWFSC bottom trawl surveys be removed from the assessment model due to the type of habitat that is sampled, though the triennial survey was retained in at least one of the models developed in the final version of the update. The California Cooperative Oceanic Fisheries Investigations (CalCOFI) index of spawning abundance is an ichthyoplankton survey and therefore samples primarily pre-recruit. All other fishery-dependent CPUE time-series have been truncated between 1996 and 2002 as a result of management changes that likely influence catch rates (MacCall, 2007a). The hook and line survey is therefore an important fishery-independent index of the bocaccio population for use in that assessment.

Selectivity for the hook and line survey is likely dome-shaped due to its use of intermediate-sized hooks and the interaction between ontogenetic shifts in depth distribution and the limited distribution of sampling locations. This means that we do not know, *a priori*, what demographic segment of the population is most vulnerable to the sampling design. This can only be reconciled within the stock assessment model via estimation of an appropriate selectivity curve. Therefore, the recent trends reported here are not directly comparable to other indices or spawning biomass trajectories reported in the most recent assessment. Indeed, a stock increasing rapidly via one or more large year classes may actually show a short-term decline in the numbers of fish over intermediate size categories. However, to the degree that the bocaccio demographic indexed here is an important contributor to stock status, the relatively flat trend appears contrary to the assessment-estimated increase over the last 5 years (Field et al., 2009).

For any survey method that has a finite capacity for captured individuals it is important to understand whether gear saturation is occurring, and over what relative levels of population abundance the survey might be expected to be proportional to change in abundance. Observed levels of gear saturation for the hook and line survey by year (Fig. 9) indicate that complete saturation (by all species captured) has never occurred. Saturation levels (as represented by the percentage of hooks deployed at a site that catch fish) of 50% or less have been observed at approximately 65% of sites sampled in any survey year. However, saturation levels of greater than 50% may not necessarily result in significant changes in the relationship between survey catch rates and local abundance. Longline studies have suggested that many species are able to locate available baits relatively quickly, even when few remain (Sigler, 2000.) In areas of very high productivity, it is probable that there are instances where the survey's catch rates are no longer proportional to local abundance. Although no site sampled during the survey's first 5 years has resulted in fewer than three unoccupied hooks, saturation levels of greater than 80% have been observed at approximately 10% of all sites sampled through 2008. The presence of unoccupied hooks at all sites, as well as the prevalence of sites with very low catch rates of bocaccio (and other species) which an expanding population could occupy, indicate that modest changes in the bocaccio population (such as rebuilding toward the current fishery management targets) may be detected in hook and line survey catch rates. The power analysis provides a very rigid test of whether future changes in the bocaccio population could be detected. By evaluating changes in a single year we simulate rapid identification of trends, when in practice assessments are only conducted biennially, and trends that are not detected in a single year may be more easily identified over the course of several survey years.

Several issues of potential concern related to the design and analysis of the hook and line survey warrant additional discussion. For example, the survey's fixed-site design raises the potential for localized depletion at survey sites as a result of annual removals from the same relatively small area of seafloor. To date, no evidence of site-specific depletion has been observed during sampling as catch rates for target species do not significantly decline with each progressive drop and cumulative removals (Harms et al., 2008). Further, given that many of the survey's fixed sites are current and former fishing "hotspots", it is likely that considerably more fishing effort has been expended at these sites in the past than under the current management system where many sites are located in depths closed to fishing or are subject to significantly more restrictive bag limits. Another concern is the potential that results from a survey conducted only at fixed sites may not be able to discern between shifts in a species' distribution and changes in its abundance. Although this is unlikely because the survey samples a variety of habitats throughout a bathymetrically and oceanographically diverse region, increasing the number of sample sites might be helpful if additional funding and charter time become available. The CCAs are not currently included in the sampling design, so the degree to which trends in those areas have diverged from the rest of the SCB will influence the relationship between this survey and total bocaccio abundance. The slightly higher catches observed in 2004 may be due to seasonal changes in distribution as a result of that survey having been conducted approximately 6 weeks later in the year than the 2005–2008 surveys.

The pros and cons of simple vs. complex analysis models depend on the intended purpose of such standardization and may differ greatly among applications. Specifically, if a survey were to be used for predictive purposes, or explicitly in a data-based harvest control rule, then the main goal might be to reduce uncertainty and minimize the probability of incorrectly identifying a trend that does not exist. However, this analysis is intended to provide fishery-independent information for use in integrated stock assessment models. Even after many simplifying assumptions have been made, there is considerable sampling uncertainty due to the limited quantity of effort that can be deployed during the course of the survey as well as uncertainty due to the many relationships that are incompletely understood (e.g., effort via fishing time, number of hooks, location of hooks in the gangion, environmental variables such as current, wind, moon phase, etc.). Unfortunately, as with all fisheries

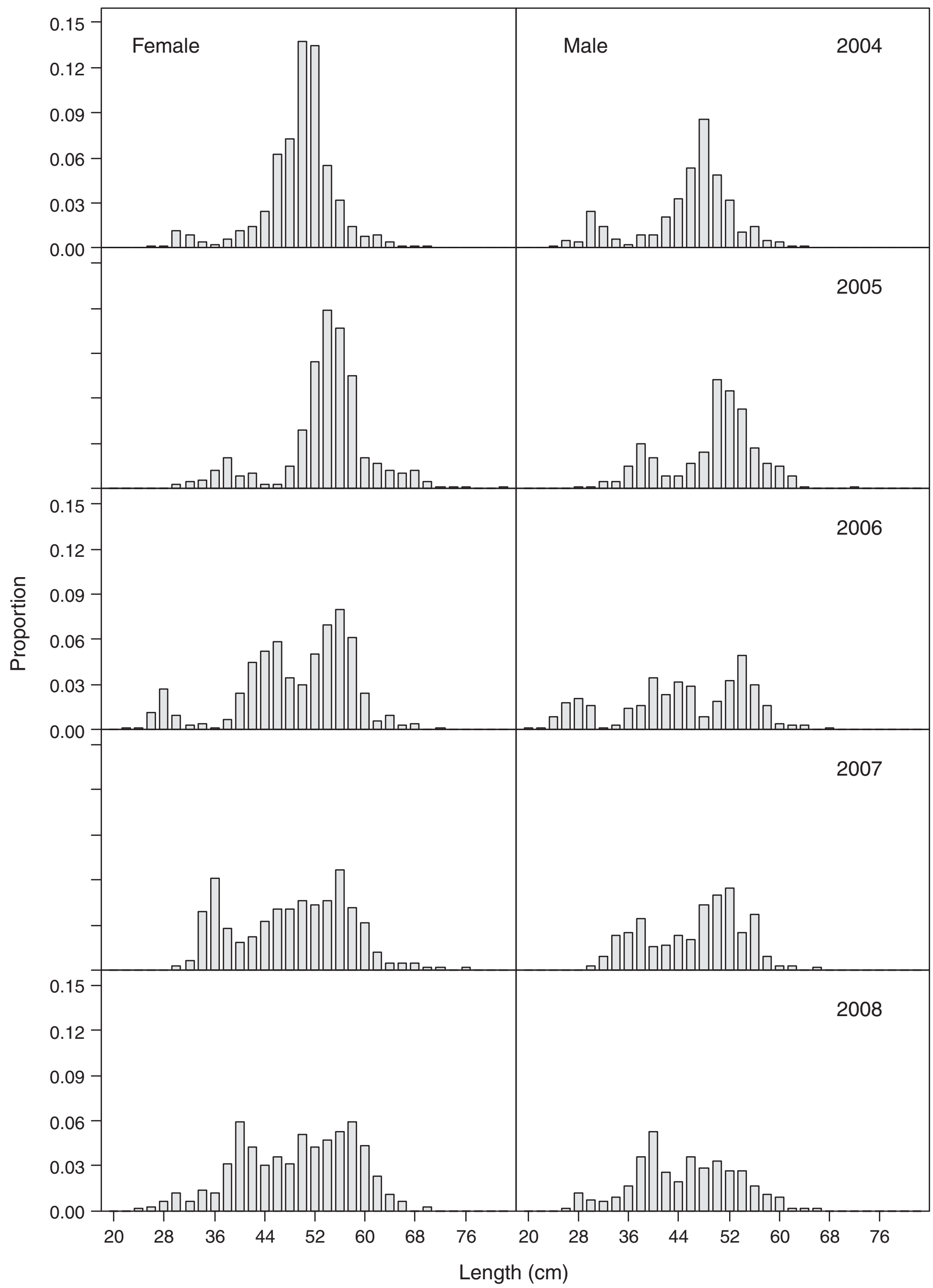

Fig. 8. Annual length frequency distributions for female (left column) and male (right column) bocaccio caught during the hook and line survey, 2004–2008. Unsexed specimens accounted for less than 1% of all specimens in any year and were excluded. Female and male specimens sum to 100% in each year to show annual proportions.

data, even a highly standardized survey design is subject to many events that are very undesirable for a pure experimentally replicated design and this can greatly increase the variance of the results and introduce substantial bias if unaccounted for, particularly in short time-series. The goals of this analysis are to better understand these relationships and to generate a time-series of relative abundance that accurately reflects these sources of uncertainty in such a way that they can be propagated through to assessment results where, often, little other information is available. As additional data are collected and the hook-and-line time-series grows longer, the underlying relationships governing covariates and the inherent site-level differences (as well as the resultant index) will become more precisely estimated.

We are exploring the use of simulation testing for development of optimal site allocation schemes for future iterations of the survey. For example, re-visiting sites that have been sampled in only 1 or 2 years would not only provide information for estimating the current year's abundance, it can contribute information to the estimates (and precision) for previous years by providing additional contrast in year to year changes in abundance. Investigation of whether bias and variance estimates were sensitive to the inclusion of sites sampled in only 2 or 3 years revealed consistency in the estimated trends, but the topic will be revisited as more data are collected.

The methods described here are applicable for developing abundance indices for several other species of rockfish in the region

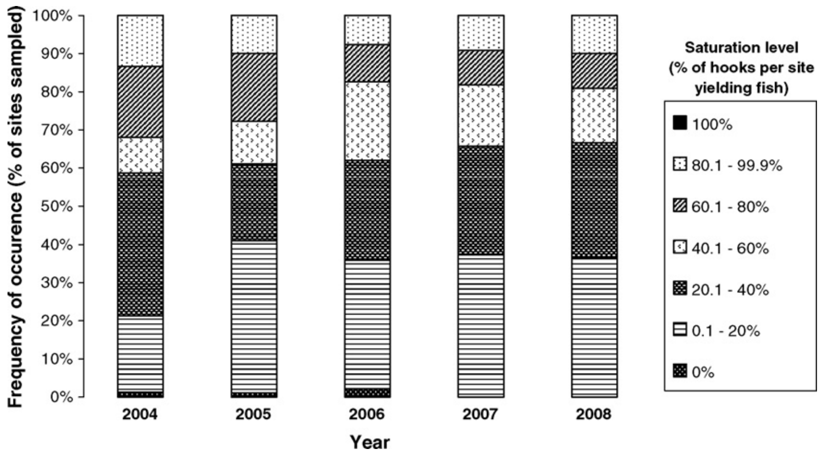

Fig. 9. Frequency of gear saturation as a percentage of hooks deployed per site that catch a fish, by year.

including vermilion rockfish (*S. miniatus*),[2] greenspotted rockfish (*S. chlorostictus*), speckled rockfish (*S. ovalis*), and starry rockfish (*S. constellatus*). These species' highest abundances occur within the SCB, are commonly encountered during the hook and line survey, and are subject to the same fishery-independent data limitations as bocaccio. Although the general method developed here is directly applicable, the process of variable selection will need to be repeated for each species. Further, a model-based index for any species using this approach must be recalculated as each new year of data is added, updating the results of the entire time-series.

The hook and line survey complements existing NWFSC acoustic and trawl surveys by providing fishery-independent observations of groundfish in areas of rocky, generally untrawlable habitats. The bottom trawl survey, a collaboration between agency scientists and the commercial fishing industry, has benefited from the knowledge and skills of its chartered vessel captains who have decades of firsthand experience with the gear, fishing grounds, and techniques associated with trawl fishing (Turk et al., 2001). Similarly, the hook and line survey has been improved through its close collaboration with the local sportfishing fleet. In addition to the direct benefits gained by the exchange of ideas and information with the fishermen, cooperative research with the industry is encouraged by statute through the Magnuson–Stevens Act, and research conducted collaboratively is more likely to be perceived by industry as credible and hence may result in increased compliance with subsequent management decisions (Rice and Richards, 1996).

Acknowledgements

The NWFSC is grateful for the 6 years of hard work and dedication of Capt. Mike Thompson and the crew of the F/V Aggressor and Capt. Joe Villareal and the crew of the F/V Mirage. Thanks are also due to the captains and crews of the F/Vs Amigo, Nikki J, and Sea Horse as well as many other sport and commercial fishermen who contributed fishing locations, advice, and expertise during the development of this survey. We would also like to recognize the following individuals for their participation in the design, conduct, logistics, or analysis of the Southern California shelf rockfish hook and line survey: Jim Benante, Matt Barnhart, Tim Athens, Keith Bosley, Andy Brooks, Toby Carpenter, Elizabeth Clarke, Eli Coplen, Al Cramer, Jen Cramer, E.J. Dick, John Field, Bob Fletcher, Ken Franke, Kate Guthrie, Pete Halmay, Owen Hamel, Don Hansen, Chris Harvey, Tom Helser, Jennifer Hempelmann, Beth Horness, Melanie Johnson, Aimee Keller, Dave Kinley, John LaFargue, Bob Lauth, Phil Levin, Alec MacCall, Scott McEntire, Jon McVeigh, Kevin Morgan, David Ono, Pat Patterson, Sylvia Pauly, Don Pearson, Kevin Piner, Steve Ralston, John Richards, Gerry Richter, Keith Sakuma, Victor Simon, Ben Susman, Stan Tomich, Waldo Wakefield, and Curt Whitmire. This research has been conducted in conjunction with federal Scientific Research Permit nos. SRP-05-2003, SRP-03-2004, SRP-01-2005, SRP-01-2006, and SRP-01-2007; California Department of Fish and Game permit nos. SC-006780, SC-006781, SC-006782, and SC-008009; and Channel Islands National Marine Sanctuary Permit nos. CINMS-2007-002 and CINMS-2008-006. Three anonymous reviewers provided thoughtful suggestions which improved the manuscript and clarity of the analysis.

[2] The recent delineation of the sunset rockfish (*S. crocotulus*), a cryptic form of vermilion rockfish (Hyde et al., 2008) provides additional complexity for research and management (MacCall, 2005). However, tissue samples from all specimens caught during the hook and line survey have been retained for genetic analyses and may provide the ability to generate separate indices for both species.

References

Burnham, K.P., Anderson, D.R., 2002. Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach, second ed. Springer-Verlag, New York.

Field, J.C., Dick, E.J., Pearson, D., MacCall, A.D., 2009. Status of bocaccio, *Sebastes paucispinis*, in the Conception, Monterey, and Eureka INPFC areas for 2009. In: Status of the Pacific Coast Groundfish Fishery through 2009, Stock Assessment and Fishery Evaluation: Stock Assessments and Rebuilding Analyses. Pacific Fishery Management Council, 7700 NE Ambassador Place, Suite 101, Portland, OR 97220-1384, 254 pp.

Harms, J.H., Benante, J.A., Barnhart, R.M., 2008. The 2004–2007 hook and line survey of shelf rockfish in the Southern California Bight: estimates of distribution, abundance, and length composition. U.S. Dept. Commer., NOAA Tech. Memo. NMFS-NWFSC-95, 110 pp.

Helser, T.E., Stewart, I.J., Whitmire, C.E., Horness, B.H., 2007. Model-based estimates of abundance for 11 species from the NMFS slope surveys. U.S. Dept. Commer., NOAA Tech. Memo. NMFS-NWFSC-82, 145 pp.

Hyde, J.R., Kimbrell, C.A., Budrick, J.E., Lynn, E.A., Vetter, R.D., 2008. Cryptic speciation in the vermilion rockfish (*Sebastes miniatus*) and the role of bathymetry in the speciation process. Mol. Ecol. 17, 1122–1136.

Jagielo, T., Hoffmann, A., Tagart, J., Zimmermann, M., 2003. Demersal groundfish densities in trawlable and untrawlable habitats off Washington: implications for the estimation of habitat bias in trawl surveys. Fish. Bull. 101, 545–565.

Keller, A.A., Horness, B.H., Tuttle, V.J., Wallace, J.R., Simon, V.H., Fruh, E.L., Bosley, K.L., Kamikawa, D.J., 2006. The 2002 U.S. West Coast upper continental slope trawl survey of groundfish resources off Washington, Oregon, and California: estimates of distribution, abundance, and length composition. U.S. Dept. Commer., NOAA Tech. Memo. NMFS-NWFSC-75, 189 pp.

Keller, A.A., Simon, V.H., Horness, B.H., Wallace, J.R., Tuttle, V.J., Fruh, E.L., Bosley, K.L., Kamikawa, D.J., Buchanan, J.C., 2007. The 2003 U.S. West Coast bottom trawl survey of groundfish resources off Washington, Oregon, and California: estimates of distribution, abundance, and length composition. U.S. Dept. Commer., NOAA Tech. Memo. NMFS-NWFSC-86, 130 pp.

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

MacCall, A.D., 2005. Assessment of vermilion rockfish in southern and northern California. In: Status of the Pacific Coast Groundfish Fishery through 2005, Stock

Assessment and Fishery Evaluation: Stock Assessments and Rebuilding Analyses. Pacific Fishery Management Council, 7700 NE Ambassador Place, Suite 101, Portland, OR 97220-1384, 128 pp.

MacCall, A.D., 2007a. Status of bocaccio off California in 2007. In: Status of the Pacific Coast Groundfish Fishery through 2008, Stock Assessment and Fishery Evaluation: Stock Assessments, STAR Panel Reports, and Rebuilding Analyses. Pacific Fishery Management Council, 7700 NE Ambassador Place, Suite 101, Portland, OR 97220-1384, 57 pp.

MacCall, A.D., 2007b. Bocaccio rebuilding analysis for bocaccio 2007. In: Status of the Pacific Coast Groundfish Fishery through 2008, Stock Assessment and Fishery Evaluation: Stock Assessments, STAR Panel Reports, and Rebuilding Analyses. Pacific Fishery Management Council, 7700 NE Ambassador Place, Suite 101, Portland, OR 97220-1384, 12 pp.

Maunder, M.N., Punt, A.E., 2004. Standardizing catch and effort data: a review of recent approaches. Fish. Res. 70, 141–159.

MCMCpack, 2009. Lead developers: Andrew D. Martin and Kevin M. Quinn. http://mcmcpack.wustl.edu/wiki/index.php/Main_Page.

Pacific Fishery Management Council, 2006. Research and Data Needs 2006–2008. Pacific Fishery Management Council, 7700 NE Ambassador Place, Suite 101, Portland, OR 97220-1384, 45 pp.

R Project for Statistical Computing, 2009. http://www.r-project.org/.

Ressler, P.H., Holmes, J.A., Fleischer, G.W., Thomas, R.E., Cooke, K.C., 2007. Pacific hake, *Merluccius productus*, autoecology: a timely review. Mar. Fish. Rev. 69, 1–24.

Rice, J.C., Richards, L.J., 1996. A framework for reducing implementation uncertainty in fisheries management. N. Am. J. Fish. Manage. 16, 488–494.

Schwarz, G., 1978. Estimating the dimension of a model. Ann. Stat. 6, 461–464.

Southeast Data, Assessment and Review (SEDAR), 2007. SEDAR Grouper Assessment Review: SEDAR Supplement 1, Final Report, June 2007. South Atlantic Fishery Management Council, 4055 Faber Place, Suite 201, North Charleston, SC 29405, 174 pp.

Seger, J., 2000. Research and Data Needs 2000–2002. Pacific Fishery Management Council, 7700 NE Ambassador Place, Suite 101, Portland, OR 97220-1384, 25 pp.

Sigler, M.K., 2000. Abundance estimation and capture of sablefish (*Anoplopoma fimbria*) by longline gear. Can. J. Fish. Aquat. Sci. 57, 1270–1283.

Spiegelhalter, D.J., Best, N.G., Carlin, B.P., van der Linde, A., 2002. Bayesian measures of model complexity and fit (with discussion). J. R. Stat. Soc. Ser. B: Stat. Methodol. 64, 583–639.

Stoner, A.W., Ryer, C.H., Parker, S.J., Auster, P.J., Wakefield, W.W., 2008. Evaluating the role of fish behavior in surveys conducted with underwater vehicles. Can. J. Fish. Aquat. Sci. 65, 1230–1243.

Turk, T.A., Builder, T., West, C.W., Kamikawa, D.J., Wallace, J.R., Methot, R.D., Bailey, A.R., Bosley, K.L., Cook, A.J., Fruh, E.L., Horness, B.H., Piner, K., Sanborn, H.R., Wakefield, W.W., 2001. The 1998 Northwest Fisheries Science Center Pacific West Coast upper continental slope trawl survey of groundfish resources off Washington, Oregon, and California: estimates of distribution, abundance, and length composition. U.S. Dept. Commer., NOAA Tech. Memo. NMFS-NWFSC-50. 121 pp.

US General Accounting Office, 2004. Pacific Groundfish: Continued Efforts Needed to Improve Reliability of Stock Assessments. US General Accounting Office, Washington, DC, Report to Congressional Requesters, GAO-04-606.

Wilkins, M.E., Zimmermann, M., Weinberg, K.L., 1998. The 1995 Pacific west coast bottom trawl survey of groundfish resources: estimates of distribution, abundance, and length and age composition. U.S. Dept. Commer., NOAA Tech. Memo. NMFS-AFSC-89, 138 pp. plus Appendices.

Xiao, Y., 2004. Use of individual types of fishing effort in analyzing catch and effort data by use of a generalized linear model. Fish. Res. 70, 311–318.

Yoklavich, M.M., Love, M.S., Forney, K.A., 2007. A fishery-independent assessment of an overfished rockfish stock, cowcod (*Sebastes levis*), using direct observations from an occupied submersible. Can. J. Fish. Aquat. Sci. 64, 1795–1804.