




# Quantifying the statistical power of monitoring programs for marine protected areas

NICHOLAS R. PERKINS <sup>1,2,3,7</sup> MICHAEL PRALL,<sup>2</sup> AVISHEK CHAKRABORTY,<sup>4</sup> J. WILSON WHITE <sup>5</sup>,  
 MARISSA L. BASKETT <sup>6</sup> AND AND STEVEN G. MORGAN<sup>6</sup>

<sup>1</sup>Coastal and Marine Sciences Institute, University of California, Davis, California 95616 USA

<sup>2</sup>California Department of Fish and Wildlife, Marine Region, Eureka, California 95501 USA

<sup>3</sup>Institute of Marine and Antarctic Studies, University of Tasmania, Taroona, Tasmania 7053 Australia

<sup>4</sup>Department of Mathematical Sciences, University of Arkansas, Fayetteville, Arkansas 72701 USA

<sup>5</sup>Department of Fisheries and Wildlife, Coastal Oregon Marine Experiment Station, Oregon State University, Newport, Oregon 97365 USA

<sup>6</sup>Department of Environmental Science & Policy, University of California, Davis, California 95616 USA

*Citation:* Perkins, N. R., M. Prall, A. Chakraborty, J. W. White, M. L. Baskett, and S. G. Morgan. 2021. Quantifying the statistical power of monitoring programs for marine protected areas. *Ecological Applications* 31(1):e02215. 10.1002/eap.2215

**Abstract.** Marine Protected Areas (MPAs) are increasingly established globally as a spatial management tool to aid in conservation and fisheries management objectives. Assessing whether MPAs are having the desired effects on populations requires effective monitoring programs. A cornerstone of an effective monitoring program is an assessment of the statistical power of sampling designs to detect changes when they occur. We present a novel approach to power assessment that combines spatial point process models, integral projection models (IPMs) and sampling simulations to assess the power of different sample designs across a network of MPAs. We focus on the use of remotely operated vehicle (ROV) video cameras as the sampling method, though the results could be extended to other sampling methods. We use empirical data from baseline surveys of an example indicator fish species across three MPAs in California, USA as a case study. Spatial models simulated time series of spatial distributions across sites that accounted for the effects of environmental covariates, while IPMs simulated expected trends over time in abundances and sizes of fish. We tested the power of different levels of sampling effort (i.e., the number of 500-m ROV transects) and temporal replication (every 1–3 yr) to detect expected post-MPA changes in fish abundance and biomass. We found that changes in biomass are detectable earlier than changes in abundance. We also found that detectability of MPA effects was higher in sites with higher initial densities. Increasing the sampling effort had a greater effect than increasing sampling frequency on the time taken to achieve high power. High power was best achieved by combining data from multiple sites. Our approach provides a powerful tool to explore the interaction between sampling effort, spatial distributions, population dynamics, and metrics for detecting change in previously fished populations.

*Key words:* integral projection model; monitoring design; remotely operated vehicle; simulation; spatial point process.

## INTRODUCTION

Marine Protected Areas (MPAs), areas of restricted human use, are being increasingly established globally as a spatial management tool to aid in conservation and fisheries management objectives (Klein et al. 2015). Assessing whether MPAs are having the desired performance effects requires establishment of effective monitoring programs before, during, and after MPA implementation. Effective design and execution of monitoring programs rely on collecting data that are capable

of tracking trends in the dynamics of populations of interest. Monitoring data collected within a rigorously designed program also enable comparisons with predicted outcomes, possibly resulting in management adjustments under an adaptive management framework (Nichols and Williams 2006, Lyons et al. 2008). In this respect, a crucial component of an effective monitoring program is ensuring that sampling designs provide sufficient statistical power to detect change (Urquhart et al. 1993, Lindenmayer and Likens 2010, Urquhart 2012). Examining the relationship between sampling effort and the power achievable with different designs is especially important in the marine environment, where surveys are often technically challenging and expensive. When a network of MPAs is the focus of monitoring efforts,

Manuscript received 8 July 2019; revised 15 April 2020; accepted 19 June 2020. Corresponding Editor: Mark Henderson.

<sup>7</sup>E-mail: Nicholas.Perkins@utas.edu.au

budgetary constraints will require that researchers choose between available tools, levels of within-site sampling and the amount of spatial and temporal replication achievable across multiple sites.

A critical aspect of monitoring program design is the selection of indicators and metrics to be used in assessing program goals. In order to track trends, indicators must respond to the process of interest (e.g., changes in population size structure and abundance after fishing ceases) and be quantifiable with accuracy and precision within the limits of the survey tools being used (Skalski 2012, Hayes et al. 2015). For many biological monitoring programs, particular indicator species are chosen for this purpose because knowledge of life history traits and important environmental drivers at the species level make likely responses more predictable (Butler et al. 2012). In the case of the MPAs, the pre-implementation level of fishing mortality largely determines the expected level of recovery (White et al. 2013, Nickols et al. 2019). For example, Kaplan et al. (2019) show how expected rates of recovery differ among previously fished species in a network of MPAs and depend on life history traits and levels of pre-MPA fishing mortality. Changes in biomass will likely show a greater magnitude of response than abundance due to the expected increase in the proportion of larger individuals in MPAs (Kaplan et al. 2019). Increases in abundance and biomass of previously fished species have been documented by a number of long-term empirical studies (e.g., Claudet et al. 2006, Guidetti 2006, Barrett et al. 2007, Lester et al. 2009), but the rate and magnitude of change reported is variable, likely reflecting life history traits of the species concerned, differing levels of pre-implementation fishing mortality and inherent environmental variability.

The ability to detect population changes with high statistical power will inevitably decrease with variability in the trend of interest, which can arise from multiple sources. Environmentally driven stochasticity, such as recruitment variability dependent on oceanographic conditions (reviewed by White et al. 2019), is irreducible but quantifiable (Regan et al. 2002). Noise from measurement or sampling error depends on sampling design and therefore is both quantifiable and reducible (Regan et al. 2002), depending on budgetary constraints. Assessment of sample sizes achievable for given indicator species and associated metrics can aid in determining both their suitability and likely timeframes to detect and report on change. For example, Starr et al. (2015) assessed fisheries-independent hook-and-line surveys across MPAs and found that detecting responses in abundance and sizes of individual species was likely to take at least 20 yr. When analysis in the early stages of monitoring shows that sample sizes may not be achievable for individual species, alternative metrics can also be explored. For example, Caselle et al. (2015) analyzed monitoring data collected by divers over the decade after MPAs were established and found that it was necessary to pool the biomass of all fished species to detect MPA

effects (but see also Caselle and Cabral [2018] for an updated analysis). Any assessment of sampling effort needs to account for the sample sizes and accompanying covariate information that available survey tools are able to capture.

MPA monitoring programs are increasingly using visual survey technologies, such as remotely operated vehicles (ROVs), due to their ability to sample non-destructively and collect data over large spatial scales (e.g., Karpov et al. 2012, Haggarty et al. 2016, Huvenne et al. 2016). In particular, the ability to conduct surveys beyond scuba diving limits (typically 20–30 m) has meant that a more complete picture of the status of many species within MPAs is achievable, as the depth distribution of most species extends beyond safe scuba diving depths. Furthermore, larger ROVs are capable of collecting over a kilometer of transect data per hour, which is much more than is routinely collected in scuba surveys, affording the potential for more accurate and precise estimation of the abundance and biomass of species across the scale of an MPA. Even so, compared to the spatial scale of a typical coastal MPA (~5–50 km<sup>2</sup>), sampling is relatively sparse, and an assessment of the ability of sampling designs to detect likely levels of change is necessary.

Here, we explore the ability of differing levels of sampling effort with a ROV to detect expected changes due to MPA implementation by developing a method to combine spatial point process models, integral projection models (IPMs; Easterling et al. 2000) and simulation-based approaches. We used baseline survey data for an example species, brown rockfish (*Sebastes auriculatus*), across three MPAs and associated reference sites in California, USA as the basis for creating simulated time series of responses to MPA establishment. Spatial point process models that quantify covariate effects and spatial patterns in the distribution of our target species simulated distributions across sites. IPMs simulated changing abundance and size structure of the populations within sites due to MPA establishment, capturing the expected initial population response to MPA implementation of a “filling in” of the size structure of previously fished size classes (White et al. 2013, Kaplan et al. 2019). Using these simulated data sets, we conducted simulated ROV transect-based surveys with varying sampling effort (number of transects) in order to characterize the data captured regarding changing size and abundance of populations of brown rockfish. We then examined the statistical power of different levels of sampling effort in both space and time to detect changes in abundance and biomass. Our novel approach (1) quantifies the effect of ROV sample design and effort on the power to detect likely changes in abundance and biomass due to MPA implementation, (2) establishes likely timeframes to detect changes between MPA and reference sites given realistically achievable levels of sampling, and (3) explores the trade-offs in replication of effort across sites and the frequency of sampling.

## METHODS

### *Study system*

California recently established a state-wide network of 124 MPAs encompassing a total of 16% of state waters. Various levels of protection were afforded through a combination of state marine reserves, state marine conservation areas, state marine parks, recreational management areas, and special closures. This network was designed and established under the legislative mandate of the *Marine Life Protection Act* (MLPA) and the MPAs were implemented between 2007 and 2013. The MLPA legislation has the aim of preserving the structure and function of marine ecosystems and protection and conservation of the populations within them. There are also requirements in the legislation that the best available science is used and that the network is monitored through time and subject to adaptive management (Botsford et al. 2014). At present, the monitoring program is transitioning from baseline quantification into the long-term monitoring phase (CDFW and OPC 2018). It is therefore timely that an assessment of the scientific knowledge gained to date be made, with a view to making recommendations for the ongoing monitoring of the network. This includes choosing appropriate indicators, site selection, sampling design and the temporal replication of effort.

California's MPA Monitoring Action Plan (CDFW and OPC 2018) identified brown rockfish (*Sebastes auriculatus*) as one of a number of potential indicator species that has been the target of historical fishing effort. Brown rockfish are distributed from southern Baja California to Prince William Sound in the northern Gulf of Alaska, but are most abundant in the central and southern parts of Puget Sound and from southern Baja California to Bodega Bay, California (Love et al. 2002). They are most commonly found in shallow (<120 m) inshore reefs and show high site fidelity. They are viviparous and produce planktonic larvae that metamorphose into pelagic juveniles before settling back to rocky reef habitat. Adults grow to a maximum size of approximately 56 cm and an age of 34 yr (Shanks and Eckert 2005). Both commercial and recreational fisheries target brown rockfish. Given the estimated level of fishing effort for this species, population models indicate that it is likely to provide a good indicator for MPA performance with relatively high projected rates of increases in abundance and biomass (Kaplan et al. 2019).

### *Data collection and survey sites*

The empirical data used in this study were collected by ROV as part of baseline characterization surveys of California's MPA network between 2014 and 2016. We focused on three study areas, each consisting of a paired MPA site and associated reference site. The three areas were Bodega Head, Año Nuevo, and Montara (Fig. 1). All study areas contained populations of brown

rockfish. Variable densities of the focal species were observed across the study areas in the survey data, which allows for the examination of the interaction between abundance and sampling effort.

The sampling design for ROV surveys utilized a defined survey rectangle that covered the depth profile of rocky reef at each site (Fig. 1). The rectangles were always 500 m wide but varied in length depending on the local extent of reef. Benthic transects were conducted across the width of the rectangle (i.e., 500 m long transects) utilizing a random systematic design. A random starting point was chosen at the shallow end of the rectangle that allows the required number of equally spaced transects to be deployed across the rectangle. Typically, the aim was to acquire at least 4 km of linear transect per site. Subsequent scoring of the video transects was conducted by trained experts in the laboratory.

### *Bathymetric covariates*

We utilized multibeam maps of the bathymetry of each study area for modeling covariate associations in the empirical data and subsequent simulations. Bathymetric data were available at a 2-m resolution (i.e., each cell is 4 m<sup>2</sup>) for all sites and were acquired from the United States Geological Survey (USGS) repository (USGS 2019). A seafloor character layer developed by USGS enabling each cell to be classed as hard, mixed or soft substrate was also used. We derived all other covariates by subsequent post-processing of the bathymetric layers in ESRI ArcGIS v 10.3.1 software. The benthic terrain modeler (BTM) and terrain attribute selection for spatial ecology (TASSE; Lecours 2015) tools were used to extract the multiple covariates. Bathymetric Position Index (BPI), a measure that categorizes benthic terrain into peaks and troughs was calculated at four scales: BPI\_2\_10, BPI\_3\_15, BPI\_5\_25, and BPI\_10\_50, where the numbers represent the inner and outer radii used (in m). Vector Ruggedness Measure, a measure of benthic terrain rugosity, was calculated at two scales: VRM\_5 and VRM\_15, where the numbers represent the radius used (in m). We also calculated curvature, slope (using the TASSE tool), relative difference to mean value (RDMV from the TASSE tool, a unitless measure of relative topographic position), Euclidian distance to the nearest hard substrate, and eastness and northness, unitless sin- and cosine-transformed measures of orientation or aspect of the cell. Both depth and depth<sup>2</sup> were tested as model parameters, as depth<sup>2</sup> may capture nonlinear effects of depth, such as when a species has preference for intermediate depths. We found these terms to be highly collinear and used depth<sup>2</sup> due to the ability to capture associations at the extreme shallow and deep ends of the surveys.

### *Modeling and simulation approach: an overview*

The modeling and simulation approach consisted of five steps: (1) spatial point process modeling of the

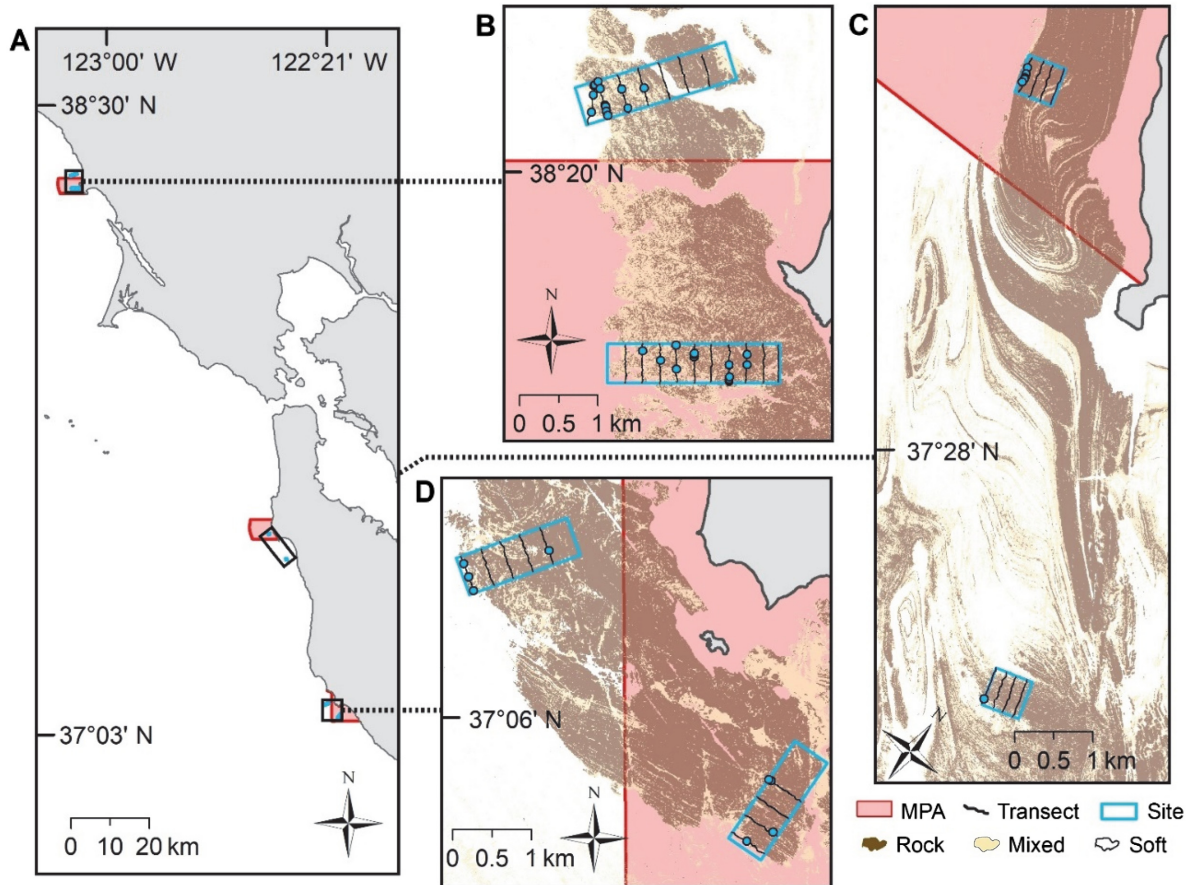


FIG. 1. (A) Map showing the study region in north-central California, USA and individual study areas: (B) Montara, (C) Año Nuevo, and (D) Bodega Head. Red shading shows the boundaries of the protected areas. Blue rectangles show the boundaries of the delineated remotely operated vehicle (ROV) sampling sites. Black lines show the navigated ROV transect lines. Blue dots show the observed locations of brown rockfish. Note that some blue dots represent aggregations of multiple brown rockfish.

baseline survey data to parameterize the spatial simulations; (2) projection of time series of abundance and size distributions of fish populations from the IPM; (3) forecast simulation of time series of fish populations across sites including explicit positions and sizes at each time step; (4) simulation of specified number of ROV transects with randomized starting points at each time step to capture sample data; and (5) analysis of the forecasted time series of estimated abundance and biomass from simulated samples (Fig. 2). Below, we outline the details of each of these processes. Example code and data are provided online (see *Data Availability*).

*Spatial point process model*

Spatial point process models are a class of models that provide a probabilistic description of the distribution of individuals (viewed as points) across a spatial domain, both in relation to where they are in terms of important covariates, such as habitat or depth, and where they are in relation to each other (e.g., Diggle 1983, Moller et al. 1998). Spatial autocorrelation in the data is explicitly

modeled, which capture patterns in distributions unaccounted for by covariates, such as might be observed with species that display aggregating behavior. Also, “marked” point processes can assign additional information to points in these models, such as the size of individuals. These models therefore provide a convenient means to model biological distributions and are being increasingly used in this setting (e.g., Niemi and Fernández 2010, Yuan et al. 2016).

To characterize the distribution of brown rockfish across each study area, we used a class of spatial point process model known as Log Gaussian Cox Process (LGCP; Moller et al. 1998) models. These models envision a continuous latent intensity surface  $\lambda(s)$  at points  $s$  across a contiguous two-dimensional domain  $D$ . However, a Poisson process likelihood, written at the point level, involves the integral  $\int \lambda(s) ds$ , which at the point level results in a likelihood  $D$

function with infinite parameters  $\{\lambda(s):s \in D\}$ . A common approach to address the problem is to assume that the  $\lambda(\cdot)$  surface is piecewise constant where each piece is small enough so that the variation of  $\lambda(\cdot)$  within that piece is not

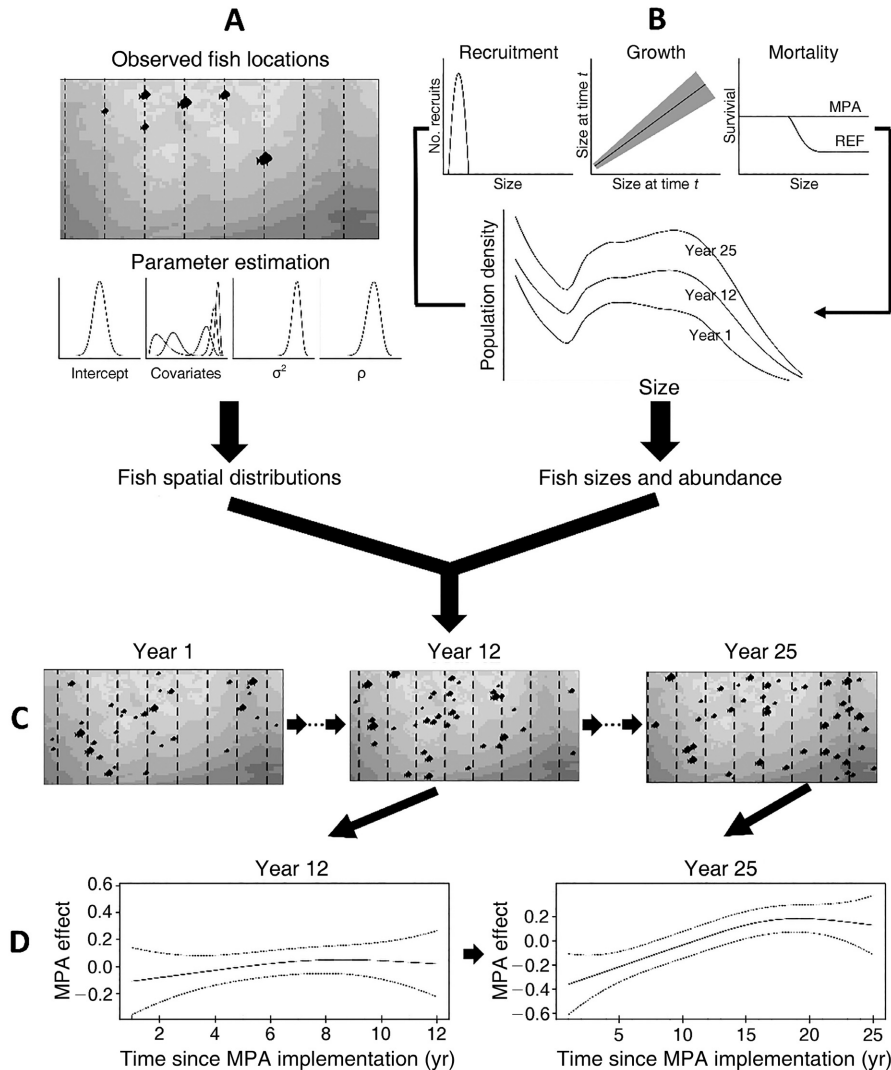


FIG. 2. Overview of the steps taken in the modeling and simulation approach: (A) Spatial point process model parameters (defined in Table 1) were estimated from empirical data and used to simulate fish spatial distributions; (B) single time step of the integral projection model (Eqs. 2–6) simulated the changing size structure and abundance of populations; (C) sampling simulations of ROV transects (dashed lines) that collected annual time series of data using different number of transects within each ROV site (only 3 yr shown here); and (D) simulated sampled data were analyzed at each time step from year 5 onward to test the ability to detect changes in trajectories between marine protected area sites and reference sites (two example times shown here). MPA, Marine Protected Area; REF, Reference area.

of interest (e.g., Waagpetersen 2004). We considered the bathymetric cell as a natural choice for this as cells are sufficiently fine scale, and bathymetric covariates are also recorded at this scale. For any such bathymetric cell  $A$ , a high (low) value of the modeled surface  $\lambda(A)$  implies an expectation of higher (lower) concentration of fish within  $A$  and is associated with observed count  $v(A)$  through a Poisson likelihood. We incorporate the bathymetric covariates within  $A$ , denoted by  $z_1(A), z_2(A) \dots, z_p(A)$ , to model  $\lambda(A)$ , on a log scale, through effect parameters  $\beta_p$ , where  $p$  refers to the value associated with the  $p$ th covariate or effect. We also use a spatial process  $\omega(\cdot)$  to explain the similarity

between intensities in nearby cells within  $D$  and model it with a zero-mean Gaussian process (GP) prior such that for any two points  $s$  and  $s'$ ,  $\text{cov}(\omega(s), \omega(s')) = \sigma^2 \exp(-\rho \|s - s'\|)$ , where  $\sigma$  and  $\rho$  represent the scale and decay-rate parameters, respectively, for the spatial process. Thus, we can write the stochastic model as

$$v(A) \sim \text{Poisson}(\lambda(A)),$$

$$\log(\lambda(A)) = \beta_0 + \beta_1 z_1(A) + \dots + \beta_p z_p(A) + \omega(s_A) \quad (1)$$

where  $s_A$  denotes the coordinates of the center of the bathymetric cell  $A$ .

We treated the covariate effects as common for all sites across the reef within the study areas, whereas the intercept and the spatial process parameters varied between individual ROV sampling sites. As study areas typically encompass the home range that a brown rockfish may inhabit, we reasoned that preferred habitat (or other covariates such as depth) within that range would likely be preferentially inhabited (or avoided for negative associations) by individual fish. We estimated intercepts and spatial processes within sites, as it was conceivable that there were effects not captured by the covariates that influenced abundance and spatial distributions at the spatial scale of a site.

We estimated the coefficients for all bathymetric covariates as well as the site-specific intercept, spatial decay-rate  $\rho$  and spatial variance  $\sigma^2$  within a Bayesian hierarchical framework. We implemented Markov chain Monte Carlo (MCMC) algorithm via Gibbs sampling for simulation of the model parameters one-by-one from their posterior distributions conditioned on the recent most values of other parameters. Because the conditional posterior distribution of spatial decay-rate parameters does not have a standard form to simulate from, we employed Metropolis-within-Gibbs (e.g., Tierney 1994) to draw its samples. All MCMC algorithms were coded using a purpose-built algorithm in the R statistical computing language (R Core Team 2019). In order to handle the large number of cells, we used a bias-corrected predictive process approximation (Banerjee et al. 2008, Finley et al. 2009) for the log-intensity function. Once the MCMC run was complete, a subset of the original posterior samples, obtained through burn-in and thinning of the chain, was retained for subsequent use. Medians of the posterior samples served as a point estimate for that parameter. We measured significance of each parameter at a level  $(1 - \alpha)$  by whether a  $(1 - \alpha)$  probability region under the empirical posterior distribution (the credible set) excluded zero. We used an  $\alpha$  of 0.10.

Total starting abundances across each site and associated starting densities (i.e., total abundance/survey area) were predicted using the set of significant covariates across each study area. We also calculated densities as the number of fishes per 100 m<sup>2</sup> as this measure of density has been commonly employed in reporting results of previous ROV surveys.

*Integral projection model (IPM)*

An integral projection model is a discrete-time population model that follows a continuous population size distribution,  $n_t(x)$  over sizes  $x$  at each time  $t$  (Easterling et al. 2000, Ellner et al. 2016). IPMs can directly simulate changes over time in population size distributions (which is the type of data collected by ROVs) and can be parameterized by continuous, size-dependent functions. Individuals of size  $x$  survive each time step according to the probability density function  $S(x)$  and grow to size  $y$  according to the probability density function  $G(y,x)$ . For

annual larval recruitment, we assume a demographically open population (White et al. 2016) with recruit density distribution  $\psi_t(y)$ . Then the population size distribution dynamics are

$$n_{t+1}(y) = S(x)G(y, x)n_t(x)dx + \psi_t(y) \tag{2}$$

where the integration is taken over all possible sizes at time  $t$ . To ensure no plausible fish sizes were excluded all possible sizes ranged from 0 to two standard deviations above the maximum length,  $L_\infty$ . All individuals experience natural mortality at a rate  $M$  regardless of size. In reference sites, individuals experience fishing mortality at a rate  $F$  with size-selectivity  $\tau(x)$ , a normal cumulative distribution function with mean of the size at entry to the fishery  $L_{\text{fish}}$  and standard deviation  $\sigma_{L_{\text{fish}}}$ . Therefore, the annual probability of survival is

$$S(x) = e^{-M}e^{-F\tau(x)}. \tag{3}$$

We modeled growth using a normal distribution with mean of the von Bertalanffy growth over a single time step starting from size  $x$  given growth rate  $k$  and maximum size  $L_\infty$

$$g_{\text{mean}}(x) = L_\infty - (L_\infty - x)e^{-k}. \tag{4}$$

Because variation around mean size-at-age typically increases with mean size (e.g., see Cope et al. 2015), we modeled the standard deviation of size in the next time step as a coefficient of variation,  $L_{CV}$  multiplied by the mean  $g_{\text{mean}}(x)$  (keeping in mind that the IPM describes changes in the population size distribution rather than individual-scale growth trajectories, which could become less variable with age) such that

$$G(y, x) = g_{\text{mean}}(x) \frac{1}{\sqrt{2\pi(L_{CV}g_{\text{mean}}(x))^2}} e^{-\frac{(y-g_{\text{mean}}(x))^2}{2(L_{CV}g_{\text{mean}}(x))^2}}. \tag{5}$$

We modeled reproduction assuming open population dynamics, i.e., all larval recruitment arrived from external sites, because the spatial scale of study areas is large relative to the movement of adult brown rockfish but small relative to the predicted spatial scale of dispersal of the pelagic larval and juvenile brown rockfish (White et al. 2014b). Thus, the density of new larval recruits of size  $y$  arriving in year  $t$ ,  $\psi_t(y)$ , is then an input of  $R_t$  recruits with mean size  $y_0$  and standard deviation  $\sigma_R$ :

$$\psi_t(y) = \frac{R_t}{\sqrt{2\pi\sigma_R^2}} e^{-\frac{(y-y_0)^2}{2\sigma_R^2}}. \tag{6}$$

Note that although larval recruitment can be highly stochastic and the actual population trajectory in a real

MPA is affected by the specific temporal pattern of annual recruitment pulses (Nickols et al. 2019), we simulated constant recruitment as a first approximation to the expected trajectory of population increase inside an MPA. In addition, we ignored immigration or emigration of adult fish from each site.

We parameterized the IPM for brown rockfish using the same approach taken by White et al. (2016) for the congener blue rockfish (*S. mystinus*), using life history and harvest parameters (e.g., the fishing rate  $F$ ) from the published literature and in the most recent brown rockfish stock assessment (Love et al. 2002, Cope et al. 2015). All parameter values and their sources are in Appendix S1: Table S1.

For both MPA sites and reference sites, we initiated simulations with the initial size structure (at the time of MPA implementation) based on the stable size distribution from simulating the IPM for 100 yr with fishing. We then iterated the IPM to project the trajectory of recovery in abundance and the changing size structure for 25 yr post-MPA implementation. For those simulations,  $F = 0$  for MPA sites but remained the same for reference sites. These simulations were made with an arbitrary value of recruitment  $R_t$ . We then rescaled the total abundance (i.e., the integral of the size distribution) in each simulation year by dividing by the integral in year 0 (the year of MPA implementation) then multiplying by the abundance estimated empirically using the spatial point process model (see *Methods: Spatial point process model*). This procedure ensured that the simulated population dynamics had both realistic total abundances and size structure that reflected the expected changes after MPA implementation (the value of  $R_t$  is factored out by the rescaling process, so results are not sensitive to the value chosen; see White et al. (2013) S3–S5). For presentation purposes, we calculated changes in total abundance as a ratio at each annual time step compared to time zero (i.e.,  $N_t/N_0$ , where  $N_t$  is the total abundance of fish greater than the minimum size observed). For the reference sites, we assumed that the population had reached equilibrium under the constant fishing mortality, and therefore there was no change in population abundance or size structure through time. Finally, we incorporated the length of detection by ROV survey for brown rockfish based on the minimum size observed in the ROV data set,  $L_{\text{detect}}$ . We assumed the detection probability of 100% for all fish over this length.

#### *Spatial forecasting simulations*

We used the parameter estimates from the spatial models and the IPM as the basis for simulating changing fish distributions, total abundances, and size structures through time. Simulated distributions and simulated sampling designs were conducted across the bathymetric grids of each site. Bathymetric grids were convenient for this purpose as each cell contained

values for the covariates used in the spatial models. For the simulations, the coefficients of the covariates were set to the medians of their posterior distributions rather than draws from the joint posterior distribution, as the focus of the study was testing the effect of sampling effort. Using different posterior draws would have resulted in a changing intercept and hence a different starting abundance for each simulation that would have been a confounding factor in determining power to detect change.

We used the rLGCP function within the spatstat package in R (Baddeley et al. 2015) to create a random realization of the spatial distribution of the brown rockfish population in each site at each time using the parameter estimates from the spatial models. The rLGCP function creates a random realization of points (fish) across the sites given the underlying covariates and spatial structuring. By creating randomized realizations, the positions of fish are different for any given time-site combination over all simulations, thereby representing the movement of fish between surveys, constrained within the model parameters. We then assigned sizes to each fish based on the proportion in each size class estimated from the IPM for that time step, thereby creating a “marked” point process. Only the portion of the IPM size distribution greater than  $L_{\text{detect}}$  (i.e., >20 cm in length) was used for this process. The step of creating realizations of the point process was not essential for the simulations, as a Poisson realization for the subset of cells traversed by the ROV for any given survey could be generated independently by integrating over the  $\lambda(s)$  (see Eq. 1). However, generating realizations allowed for visual verification of the patterns generated, which was useful in explaining the simulation process to a non-statistical audience.

For each site and time, we simulated ROV transects utilizing the established design, but with a differing number of transects with a randomized starting point. We used a range of values (8, 12, 16, and 20) for the number of transects tested within each site to test different efforts in terms of the proportion of the survey area sampled at each site (summarized in Appendix S2: Table S1). The upper limit (20 transects) is the expected maximum that could be currently conducted in a single day. The lower limit (8 transects) was chosen as the number of transects used in the baseline surveys varied from 4 to 8 (Fig. 1). We therefore simulated an increase in sampling in order to test the potential for more effort to improve the power to detect change. While real transects are likely to rarely exactly follow the planned survey lines and are also likely to vary in width along the transect line, we assumed straight line transects with a constant width for simplicity. Conveniently, the average transect width for the ROV surveys was approximately 2 m, which corresponds to the cell widths, and therefore we assumed that a strip of cells was navigated along each transect line.

In detail, for a total of 100 simulations for each site-time combination, the simulation steps were as follows:



- (1). Create raster stacks (raster package in R; Hijmans and van Etten 2012) of all covariate layers based on the 2-m resolution cells.
- (2). If the site was an MPA site, adjust the intercept (representing the mean expected fish count per cell) from the spatial model based on the time step, thereby allowing the increased abundance through time. We achieved this by back-transforming the intercept (Eq. 1; taking the exponential), multiplying it by the relevant ratio of increase for that time step (based on the IPM projection of  $N_t/N_0$ , where  $N_t = n_t(x)dx$ ), and then taking the logarithm of this value as the “new” intercept. For reference sites, the intercept remains constant.
- (3). Multiply site level covariate coefficients that were statistically significant at the 90% credible interval level in the spatial model ( $\beta$  estimates for each covariate) with cell-level covariate values in the raster stacks and then added to the intercept from step 2 to create a single value for each cell that was a linear combination of the spatial model estimates and covariates. We then turned this into an “image” (function `im` in the R `spatstat` package), representing the intensity surface across the entire site to be used in simulating the point process.
- (4). Apply the `rLGCP` (`spatstat` package) function to the intensity surface image from step 3 and the spatial variance and range parameter estimates from the spatial model to simulate a random realization of points (fish) across the site. Cartesian coordinates of each point were stored.
- (5). Assign sizes to each simulated fish proportionally by each 1-cm size bin from the IPM for a given time, resulting in a “marked” point process.
- (6). Choose a random starting point for the transects, according to the number of transects being simulated, that allowed equal spacing of the transects across the survey area. Determine then store the set of 2-m cells that were intersected by each transect line.
- (7). Store the number of and size of all fish “sampled” within cells for each simulated transect for subsequent analysis.

#### *Analysis of the forecasted data*

We analyzed the forecasted data by modeling the time series of abundance and biomass data at each time point for MPA and reference site pairs and all three study areas in a combined hierarchical model. A spatial point process model, as described above could have been fit to each simulated time point and used to estimate the differences through time. However, the spatial models were computationally intensive, and considering the large number of simulated data sets, we considered this to be impractical for the present study. Furthermore, analyzing the simulated data with a model that was structurally identical to that used to generate the data would have provided better fits than could be expected with real

world monitoring data. We instead take a conservative approach and use a simpler analysis such as may be more commonly taken with this kind of data. Abundance data consisted of the total fish sampled within each sampling site (either MPA or reference site within a study area). We calculated biomass  $B_{i,t}$  in site  $i$ , time  $t$  given the number of observed individuals  $N_{i,t,x}$  in each one centimeter size bin with a midpoint length of  $x \in [x_{\min} - x_{\max}]$  and length–weight conversion parameters  $a$  and  $b$  (see Appendix S1: Table S1 for parameter values)

$$B_{i,t} = \sum_{x=x_{\min}}^{x_{\max}} N_{i,t,x} a x^b. \quad (7)$$

To distinguish an MPA effect by comparing MPA sites to reference sites with fishing (before MPA establishment and/or outside MPAs), we used generalized additive models (GAMs). GAMs provide the flexibility of modeling the expected nonlinear trends of abundance and biomass through time for the MPA sites (e.g., White et al. 2013, Kaplan et al. 2019). Due to limited adult movement and widespread larval movement, we did not consider “spillover” effects. Also, for simplicity, we did not consider wholesale increases in reproductive output or the possibility of displaced fishing effort outside the MPA, and so reference sites could represent either before MPA data or outside-MPA data. GAMs were fitted at each time step from year 5 onward to test for a detectable MPA effect. We modeled the expectation,  $E(\cdot)$  of the observed abundance or biomass,  $y_i$ , where  $\alpha$  is an intercept,  $m_i$  is the MPA status for the  $i$ th site,  $\beta_1$  is the MPA effect,  $f_1(t_i)$  is a smooth function of time, and  $f_2(t_i)$  is a smooth deviation from  $f_1(t_i)$  due to the MPA status of sites. For abundance GAMs the variation around the fitted expectation was Poisson and for biomass GAMs it was Tweedie (e.g., Foster and Bravington 2012)

$$\log(E(k_i)) = \alpha + \beta_1 m_i + f_1(t_i) + f_2(t_i). \quad (8)$$

A generalized additive mixed model (GAMM) was used for the time series that included all sites, where model terms were the same as for (8), but with the additional term  $u_i$ , which is a study area random effect with  $u_i \sim N(0, \delta^2)$

$$\log(E(k_i)) = \alpha + \beta_1 m_i + f_1(t_i) + f_2(t_i) + u_i. \quad (9)$$

We examined model residuals for temporal autocorrelation in a large subset of models and found no evidence of a consistent temporal lag in correlation.

We determined the statistical power to detect the forecasted level of change at each study area by calculating the proportion of simulations at each time step with a significant (at  $\alpha = 0.05$ ) difference trend  $f_2(t_i)$  in the abundance or biomass trend between the MPA(s) and the associated reference site(s). We also calculated the power at each time step/sampling effort combination. In order to assess the uncertainty associated with using 100 simulations, a



random resample with replacement of 100 of the forecasts was taken and repeated 100 times. Final results are the mean of these 100 resamples as well as the maximum and minimum power calculated over the 100 resamples.

We tested the implications of reducing the frequency of temporal revisits to the “panel” of three study areas to every two or three years. We subsetted the forecasted data to the appropriate time intervals across all simulations and analyzed the combined biomass estimates across the three study areas.

For all combinations of study area/metric/number of transects/sampling frequency, we assessed the ability of the sampling design to achieve high power (>80%) within the time tested (e.g., Urquhart et al. 1993, Urquhart 2012, Andersen et al. 2019). Sampling schemes that achieve high power in a shorter timeframe are preferred for a monitoring program but must be balanced within budgetary and time constraints. Where high power is not achieved in the tested timeframe, sampling schemes that result in a trajectory that is likely to achieve high power faster than others are preferred. While an understanding of the sampling effort required to achieve high power would be informative, we limited our simulations to the number of transects that can currently be conducted at a site within a single day of sampling (i.e., 20 transects). We did not conduct null hypothesis tests on the simulation outputs, because the results of such tests depend on sample size, which is essentially arbitrary in a simulation context (see White et al. [2014a] for additional explanation).

We quantified the precision of estimates of abundance and biomass dependent on sampling effort by calculating the coefficient of variation (CV) of estimates for the reference sites in each study area. The reference sites were chosen for this purpose as the abundance and biomass were kept constant over all simulations, whereas, for the MPA sites, abundance and biomass varied for each year simulated, which would also affect precision. We calculated the CV (the standard deviation divided by the mean) of estimated abundance and biomass for each reference site over the 2,500 simulations (100 simulations each with 25 yr) for 8, 12, 16, and 20 transects.

Finally, we conducted two additional simulation studies in order to examine (1) the implications of using the Bayesian posterior median estimates for our simulations rather than the full Bayesian posterior distribution and (2) the level of spatial aggregation when considering the full simulated point patterns across sites. Full details of the approaches taken and the results are given in Appendix S3 (implications of using posterior medians) and Appendix S4 (level of spatial aggregation in simulated distributions).

## RESULTS

### *Spatial point process and integral projection model results*

In the spatial point process models across study areas, the statistically significant ( $P < 0.10$ ) bathymetric

covariates predicting brown rockfish density were depth<sup>2</sup>, bathymetric position indexes BPI\_2\_10, BPI\_5\_25 and BPI\_10\_50, curvature, slope, vector ruggedness measures VRM\_5 and VRM\_15, eastness, northness, seafloor character of mixed, and distance to hard substrate. However, depth<sup>2</sup> was the only statistically significant covariate across all study areas (Table 1). Other significant covariate effects differed between study areas.

Predicted starting densities varied considerably across the three study areas, with Bodega Head having the highest density, Año Nuevo having an intermediate density and Montara having a low relative density, particularly at the reference site (Table 1).

The spatial effects differed across the three study areas, with the median for the spatial decay-rate parameter  $\rho$  (Table 1) indicating that the practical range (the distance at which spatial correlation is 0.05) varied from approximately 37 m at the Año Nuevo MPA site to approximately 800 m at the Bodega Head reference site. The practical range was between 150 and 600 m at the other sites. The short practical range at the Año Nuevo MPA site is related to only four observed brown rockfish, with two of these in close proximity. There is also a much larger uncertainty with the estimate for the spatial decay-rate parameter at this site (Table 1).

Median values for spatial variance  $\sigma^2$  (Table 1) also differed among sites, with a high at the Año Nuevo MPA site of 0.53, and a low at the Montara MPA site of 0.22. This parameter reflects the importance of the spatial field  $\omega(\cdot)$  in explaining the intensity in counts of fish across the cells, with smaller values indicating the covariates are mostly explaining the distribution. However, the values for  $\sigma^2$  need to be interpreted in relation to the other model parameters that determine the overall intensity of fish counts and the associated variance at each site. Therefore, direct comparison of values of  $\sigma^2$  between study areas is complicated as covariate effects, mean intensity and associated variability differ between study areas.

The IPM for brown rockfish predicted approximately a 44% increase in abundance and 95% increase in biomass after 25 yr for brown rockfish >20 cm in length (Fig. 3). These changes were due to a filling in of the initial truncated size structure of the population in the MPA sites through previous removal of legal-sized individuals (Fig. 2B).

### *Forecasted simulation results*

Statistical power to detect increases in abundance at MPA sites compared to reference sites generally increased with time and sample size (Fig. 4). As expected, higher power was achieved with greater sampling effort, with clearer differences where starting densities were higher; however, high statistical power (>80%) to detect forecasted changes in abundance between individual or aggregate MPA and reference site pairings

TABLE 1. Median posterior parameter estimates and 90% credible intervals and starting densities estimated from the spatial models used in the simulation for each of the three sites.

Bathymetric covariates	Bodega Head			Año Nuevo			Montara		
	Median	5%	95%	Median	5%	95%	Median	5%	95%
Depth <sup>2</sup>	0.25	0.06	0.39	0.21	0.05	0.558	0.82	0.21	1.09
BPI_2_10							0.28	0.02	0.43
BPI_5_25	0.16	0.12	0.26						
BPI_10_50							0.40	0.12	0.59
Curvature							-0.78	-0.94	-0.62
Slope	-0.21	-0.33	-0.10				0.37	0.24	0.61
Eastness				-0.44	-0.80	-0.20	-0.51	-0.59	-0.35
Northness	-0.07	-0.01	-0.01	0.31	0.17	0.70	-0.16	-0.25	-0.09
VRM_5	0.37	0.29	0.49						
VRM_15	-0.40	-0.75	-0.27	0.34	0.08	0.50			
SF_character_mixed	0.14	0.06	0.26						
Distance to hard	-0.26	-0.45	-0.12	-0.84	-1.16	-0.22			
Intercept MPA	-5.10	-6.24	-5.14	-7.47	-7.73	-6.80	-6.90	-7.13	-6.29
Intercept REF	-4.90	-6.50	-5.40	-7.11	-7.37	-6.07	-6.87	-7.11	-6.63
Spatial decay-rate parameter ( $\rho$ ) MPA	0.01	0.00	0.00	0.08	0.01	1.58	0.01	0.01	0.03
Spatial decay-rate parameter ( $\rho$ ) REF	0.00	0.00	0.01	0.01	0.01	0.01	0.02	0.01	0.08
Spatial variance ( $\sigma^2$ ) MPA	0.17	0.09	0.51	0.53	0.26	1.12	0.22	0.14	0.53
Spatial variance ( $\sigma^2$ ) REF	0.23	0.13	0.51	0.34	0.24	0.58	0.44	0.26	1.01
Starting density (fish/100 m <sup>2</sup> ) MPA	0.17	0.17	0.17	0.09	0.09	0.09	0.06	0.06	0.06
Starting density (fish/100 m <sup>2</sup> ) REF	0.18	0.18	0.18	0.09	0.09	0.09	0.02	0.02	0.02

Notes: For bathymetric covariates, only values for significant covariates that were used in the final models are provided. Mean starting and ending abundances were averaged over the 100 simulation runs. Starting and ending abundances for the reference site were the same. VRM is Vector Ruggedness Measure, a measure of rugosity, with the number referring to the radius (meters) used for calculation. BPI is Bathymetric Profile Index, a measure of the position in the landscape, with the numbers referring to the inner and outer radii used for calculation. SF\_character\_mixed is a categorical variable defining when the seafloor (SF) related to a bathymetric cell was classified as being mixed substrate. MPA, Marine Protected Area; REF, Reference area.

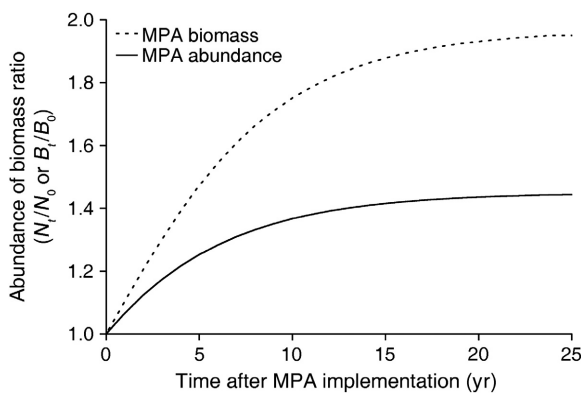


FIG. 3. Predicted trajectories for the abundance and biomass of brown rockfish in response to MPA implementation, based on the integral projection model ( $N_0$ , abundance at time 0;  $N_t$ , abundance at time  $t$ ;  $B_0$ , biomass at time 0;  $B_t$ , biomass at time  $t$ ). Ratio change (vertical axis) represents the ratio compared to the population at time of the baseline survey. The reference site was assumed to have a constant abundance and biomass through time under equilibrium condition of constant fishing pressure. The projection incorporates a length of detection of 20 cm for brown rockfish.

could not be achieved with even the highest sampling efforts tested (20 transects). The highest power achieved was for the Bodega Head study area, where mean power of approximately 50% was achieved with 20 transects

after 25 yr. Mean power to detect the simulated change in abundance at Año Nuevo never exceeded 40%, and never exceeded 25% at Montara, regardless of the level of sampling effort tested.

Higher mean power was achievable for detecting changes in biomass compared to abundance for all study areas (Fig. 5). Similar to abundance, the highest mean power to detect change in biomass was achievable at the Bodega Head study area, where a mean 80% power was achieved with 20 transects at year 25. Mean power to detect changes in biomass peaked at 73% at Año Nuevo and 47% at Montara. However, the combined study area analysis for biomass was capable of achieving high mean statistical power with all sampling efforts tested except eight transects. By combining the three study areas, high mean statistical power was achieved after 13, 16, and 18 yr with 20, 16, and 12 transects, respectively.

Reduced temporal revisit plans of every two or three years to the three study areas were able to achieve high mean statistical power to detect differences in the trajectories for biomass within the 25-yr forecast (Fig. 6). With the highest sampling effort (20 transects) high mean power took 3 yr longer with biennial visits compared to annual visits and 5 yr longer with triennial visits (Figs. 6, 5). High mean power could not be achieved with only eight transects for either revisit scheme.

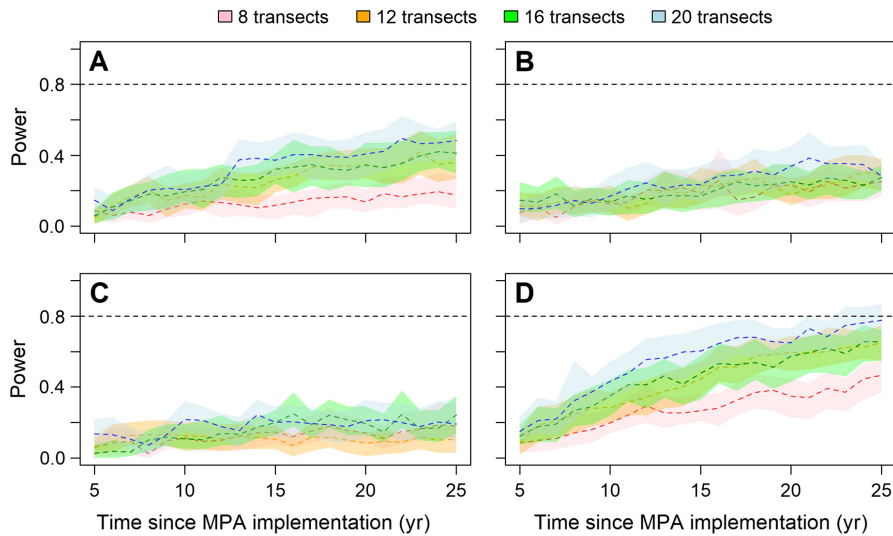


FIG. 4. Statistical power to detect forecasted changes in abundance of brown rockfish for each of the three study areas: (A) Bodega Head, (B) Año Nuevo, (C) Montara, and (D) an analysis combining all three study areas. Different colored lines and polygons represent different sampling efforts in terms of the total number of ROV transects used in the simulation. Polygons represent the distribution of the maximum and minimum power calculated at each time step by resampling the 100 forecasted data sets 100 times. The colored dashed lines represent the mean of the 100 resamples. The black dashed line indicates 80% power. Power was assessed at each year as the proportion of 100 simulations with a significant difference in trajectories between MPA and reference site populations.

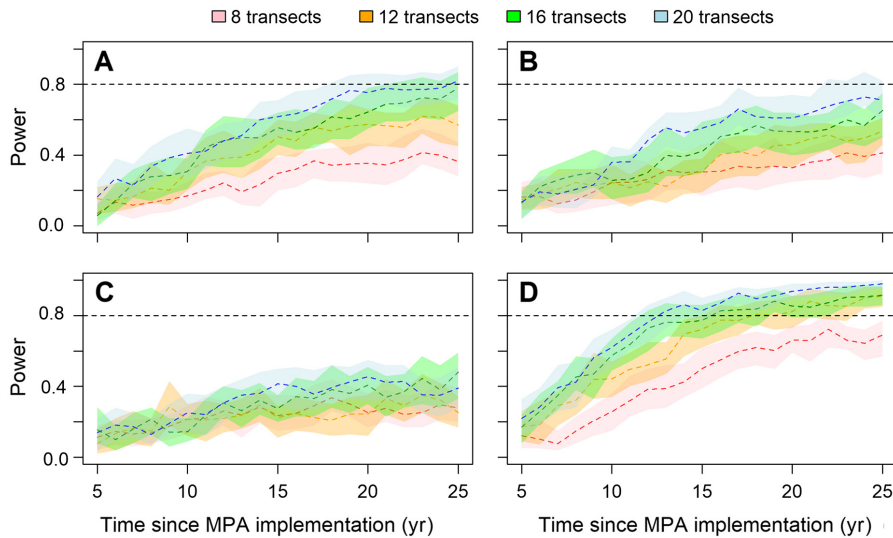


FIG. 5. Statistical power to detect forecasted changes in biomass of brown rockfish for each of the three study areas: (A) Bodega Head, (B) Año Nuevo, (C) Montara, and (D) an analysis combining all three study areas. Different colored lines and polygons represent different sampling efforts in terms of the total number of ROV transects used. Polygons represent the distribution of the maximum and minimum power calculated at each time step by resampling the 100 forecasted data sets 100 times. The colored dashed lines represent the mean of the 100 resamples. The black dashed line indicates 80% power. Power was assessed at each year as the proportion of 100 simulations with a significant difference in trajectories between MPA and reference site populations.

CVs attained for both abundance and biomass estimates for each reference site were lower (more precise) with increased sampling effort. Increased precision was also correlated with increased starting densities at each site (Fig. 7 and Table 1). Notably, the decrease in CV did not show a clear signal of flattening out at a high number of transects, indicating that even with 20

transects there were not diminishing returns to precision obtained by adding more sampling effort.

Mean sample sizes (number of predicted fish observed) achieved over the 100 simulations ranged from 5 fish at the Montara reference site with 8 transects, to 48 fish at the Bodega Head MPA site with 20 transects at year 25 (Appendix S2: Table S2). The mean sample

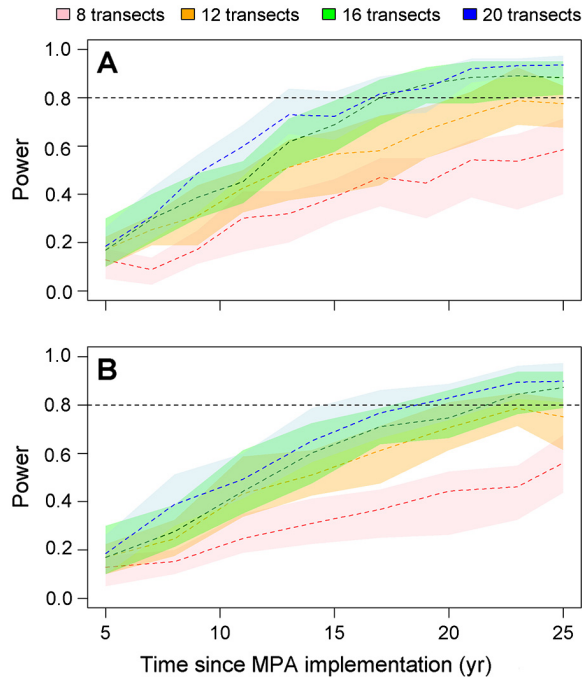


FIG. 6. Statistical power to detect changes in biomass of brown rockfish combining biomass data from the three study areas with a reduced revisit sampling design of (A) every 2 yr after year 5 and (B) every 3 yr after year 5. Shaded polygons show the range of power values obtained by resampling the 100 forecast data sets 100 times. Colored dashed lines show the mean of the 100 resamples.

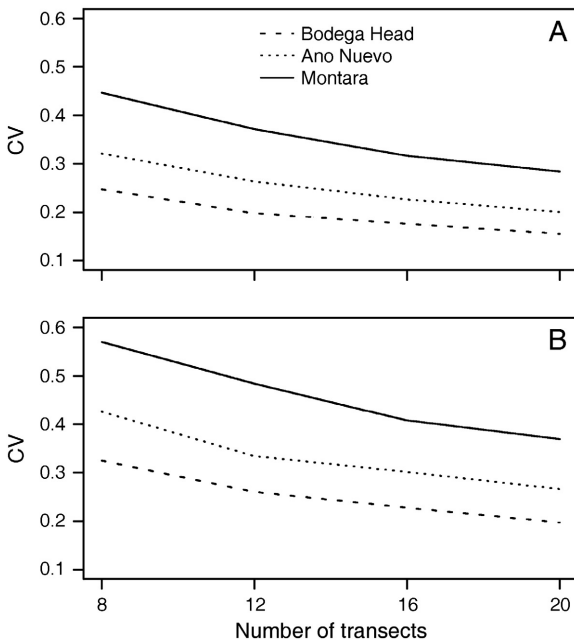


FIG. 7. Coefficient of variation (CV) for the reference sites in each study area with 8, 12, 16, and 20 transects for (A) abundance and (B) biomass. CV was calculated as standard deviation/mean over 2,500 simulations (100 simulations  $\times$  25 yr) for each sampling effort and study area combination.

sizes achieved for a given sampling effort corresponded with our expectations given the empirical data. For example, sample sizes from the empirical data with 8 transects at Bodega Head MPA (15 fish observed) and reference site (16 fish observed) both fell within one standard deviation of the simulated mean sample abundance with the same sampling effort (Appendix S2: Table S2).

Our additional simulation study examining the implications of using posterior medians (Appendix S3) demonstrates that using our simulation approach captured the major distributional patterns of brown rockfish across sites. Analysis of example mapped point patterns (i.e., complete distributions simulated across sites) showed that distributions of brown rockfish in our study did not display strong levels of aggregation (Appendix S4).

#### DISCUSSION

A power analysis before data are collected can inform choices about sampling design, effort, and length of study. However, there is considerable debate about how much power is sufficient for strong inferences (Fairweather 1991, Di Stefano 2003), and power analyses rely on inevitably uncertain predictions of future effect sizes and how those effects might be detected by a particular sampling approach. We have introduced an approach for

conducting power analysis based on predicted effects and detectability in order to inform temporal and spatial sampling design for the monitoring and adaptive management of MPAs. While we have focused on a ROV as the sampling platform, this approach could be extended to incorporate other sampling platforms (including SCUBA diver visual surveys) and other environments, although it is most suited to survey methods that record explicit spatial positions of observations.

For our study system, based on our method of combining spatial point process modeling and integral projection modeling to simulate time-series data sets across three Californian MPAs, we found the potential for detectable changes in biomass but not abundance within 13 yr. In addition to increased within-site sampling and across study area data aggregation, detectability increased with increasing starting densities across sites. Reducing sampling frequency had a smaller effect on detectability than reducing sampling intensity. These results carry an important message for managers of California's network of MPAs, and indeed MPAs elsewhere, as they (1) clarify the time frames likely to be required to report change to stakeholders and (2) explore the factors affecting the statistical power to detect likely levels of change, thereby providing guidance regarding the selection of indicators and metrics for long-term monitoring.

#### *Timeframes to achieve high power in detecting MPA effects*

Determining likely timeframes for detecting the effect of protection from fishing pressure is crucial from the management perspective of selecting appropriate indicators for MPA monitoring (see Kaplan et al. 2019), and also for informing stakeholder expectations with respect to reporting on the effectiveness of MPAs. Our simulations demonstrated that the timeframes to achieve high mean statistical power to detect change for an example species with a ROV ranged from 13 yr to over 25 yr. This was despite testing a considerable increase in sampling effort compared to baseline surveys and incorporating data from three study areas. The shortest timeframe was achieved by combining data across study areas and using biomass as the metric. For individual study areas, a minimum of 25 yr was required with the highest within-site sampling effort to detect the projected trends. Similar timeframes have also been reported in a study based on hook-and-line surveys of MPA and reference sites on the central coast of California, where it was concluded that 20 yr or more would be required to detect changes (Starr et al. 2015). In marine ecosystems, which are often highly variable in time and space, considerable "noise" is likely to exist across time series making trends harder to detect and thus reporting timeframes may be even longer than those reported here. When coupled with relatively small sample sizes due to the technical difficulties of collecting data in marine

ecosystems, long periods of data acquisition will likely be necessary (e.g., Perkins et al. 2017). Communicating this fact to stakeholders, including the general public, is a necessary component of effective management, particularly when reporting cycles are likely to be much shorter.

#### *Factors affecting the power to detect change*

A complex interplay of factors, such as within-site sampling effort, the number of study areas, starting densities, likely levels of change in chosen metrics and the frequency of revisits, all affect the ability of monitoring programs to detect changes in target populations through time. Simulation studies provide a powerful means of exploring the importance of different factors, and in particular how sampling design choices influence the length of time necessary to accumulate statistical power for detecting trends (Field et al. 2005). Therefore, it is surprising that the number of simulation studies for monitoring programs aimed at detecting changes in abundance through time is currently limited (Andersen et al. 2019). By utilizing empirical data to inform models, our simulation approach provides a means of assessing the factors influencing the power to detect change in an MPA monitoring context. Below, we discuss these factors and focus on insights concerning improved outcomes for long-term MPA monitoring.

#### *Level of within-site sampling*

Within-site sample size is known to be a key determinant of statistical power for monitoring programs (Hatch 2003), particularly where target species are rare and detection probabilities are likely to be low (Strayer 1999). We found that sampling effort strongly affected the ability to detect change in a timely manner, with the lowest level of sampling unable to detect change within the time tested, whereas increasing effort could potentially detect change within 13 yr for our example species. This agrees with other research showing that insufficient within-site sampling can have a large impact on the power to detect trends (e.g., Perkins et al. 2016, Barker et al. 2019). However, we found that increases in sampling effort in the study area with the lowest starting densities (Montara, and the reference site in particular) achieved only marginal increases in power over time, despite sites at Montara being smaller and thus the percentage area sampled with an equivalent number of transects being almost four times that of the largest sites at Bodega Head. This was particularly the case for the Montara reference site, where we found high precision in both abundance and biomass was difficult to achieve. Hence, improvements in monitoring outcomes are likely to be achieved by targeting sampling effort to filling in gaps where sufficient sample sizes are likely to be obtained for the target species.

### *Initial density of indicator species*

We found that the power achievable through time was correlated with increasing starting densities (i.e., abundance per unit area), with high power being correlated with higher starting densities. This result is supported by other simulation studies that found clear relationships between starting densities and the power to detect population trends, with increasing power achievable as starting density increases (Ficetola et al. 2018, Andersen et al. 2019). While we did not explicitly test the link between starting densities and power to detect change, we found the precision of estimates (CVs) of abundance and biomass for a given sampling effort also correlated to starting densities. These findings suggest that starting densities and the associated sample sizes achievable can be an important consideration when selecting indicator species for monitoring.

### *Incorporating data from multiple study areas*

We showed that aggregating information across multiple study areas within an MPA network can provide the best means of increasing the power to detect trends in the shortest time. In a power simulation study for terrestrial fauna, Andersen et al. (2019) found that for the tested sampling effort across space and time, the highest power to detect changes in abundance was consistently achieved by maximizing the number of sites. This outcome was found to hold true irrespective of the different starting densities tested. Similarly, Ficetola et al. (2018) in another simulation study found that high power was achievable for detecting declines in abundance with the largest number of sites tested for all but the least abundant species. Therefore, when considering panels of sites to be visited, maximizing the number of sites with higher densities of target indicators will provide a better chance of detecting change earlier.

The panel of study areas or sites that are grouped for analysis needs careful consideration, with higher power likely to be achieved when grouped sites have similar responses (Weiser et al. 2018). Our simulations assumed the same trend across the three study areas, an assumption that is unlikely to hold true, as areas probably experienced different fishing pressure before MPAs were established and likely experience different levels of recruitment. Ideally, an assessment of the larger-scale spatial correlation among MPAs across a network would allow for the selection of groups of MPAs that are more likely to show similar responses. For example, Hamilton et al. (2010) showed that grouping areas biogeographically for MPAs around the Channel Islands, California, improved statistical power to detect trends. Alternatively, different trends can also be included in hierarchical models for multiple sites (e.g., Urquhart 2012, Perkins et al. 2017), although power to detect overall trends may be low when trends among sites differ substantially (Weiser et al. 2018). Further work exploring

the influence of including a larger number of study areas with differing levels of response is warranted.

### *Effect sizes and metrics for detecting change: abundance versus biomass*

The effect size is crucial to the outcome of any power study but is often uncertain in biological systems that have considerable natural variability. Our study reinforces the expectation that biomass provides a detectable signal for change in MPAs earlier than abundance because it captures increases in the proportion of larger fish as well as population size after fishing ceases (e.g., Edgar et al. 2014, Caselle et al. 2015, Friedlander et al. 2017). Our IPM showed that expected increases in biomass were more than double that of abundance of brown rockfish in MPAs (95% compared to 44%) and was intermediate across the 19 species tested in a study by Kaplan et al. (2019), providing a representative example. Congenic rockfishes, such as blue (*S. mystinus*) and vermilion (*S. miniatus*) rockfish, had larger expected biomass responses than brown rockfish so that differences likely would be detected even sooner were other factors held constant. Kaplan et al. (2019) previously showed that increases in biomass were consistently more detectable than abundance in the expected recovery of a range of species across the network of MPAs along the coast of California. Information regarding likely response levels could be used in conjunction with our findings to select indicator-study area combinations that are likely to show larger responses.

We found that the precision for biomass estimates was lower compared to abundance estimates across all our study areas; however, the lower precision was offset over time by the larger trend, with high power achievable sooner. In order to convert lengths collected in survey data to biomass, length is raised by some exponent, typically approximately 3 (3.07 for brown rockfish), which compounds errors in both the length measurement and the original estimation of the allometric exponent. This issue is seldom addressed and warrants further investigation. We found biomass estimates had higher precision at sites with higher densities, where sample sizes were typically greater than 25 (Año Nuevo) to 48 fish (Bodega Head) with the highest sampling tested, compared to <20 fish typically sampled at Montara. Therefore, sample sizes of greater than 20 brown rockfish are likely to achieve high power in detecting expected changes in biomass over periods of 25 yr or less.

High power to detect changes in previously fished populations could also be achieved by combining data from multiple species. For example, in assessing the effects of establishing MPAs in the Channel Islands using SCUBA surveys, Caselle et al. (2015) grouped biomass of all targeted species, presumably due to the noise and lack of detectable effect when examining individual species trajectories. Hence, power is likely to be increased by grouping species for analysis, although care

would need to be taken regarding whether these species and the study areas from which they were being included would likely follow similar trajectories of recovery. In an updated analysis examining a wider data set, Caselle and Cabral (2018) showed that changes in biomass for some of the more abundant individual species can be detected with a sufficient time-series of SCUBA transect data. This aligns with our findings regarding the importance of considering initial densities and the effect this has on the precision of estimates of both abundance and biomass.

#### *Temporal revisit design*

Reducing the frequency of temporal revisits across our network of three study areas to every 2 or 3 yr delayed achieving high power by approximately three to eight years compared to sampling annually. Andersen et al. (2019) found similar results in their simulation study, with high power taking 3 and 6 yr longer with biennial and triennial sampling, respectively. Whether this trade-off is an acceptable compromise to including a larger number of study areas over time is an important management decision. While high power will take longer to achieve, reducing the temporal revisit frequency allows other study areas in a network to be visited in intervening years. Optimal designs that balance spatial and temporal replication of survey effort will likely rely on incorporating data regarding spatial and temporal variability across the network of study areas for a given indicator (e.g., Urquhart et al. 1993, Hewitt and Thrush 2007, Perkins et al. 2017). While such data are likely to be sparse in the early stages of a monitoring program, incorporating the knowledge gained from ongoing surveys allows for improvements to be made in the design of future surveys (Lindenmayer and Likens 2010).

#### *The importance of covariates and spatial distributions*

As well as the density of an indicator across an area, its spatial distribution also plays a role in the power to detect changes in populations through time, with more aggregated distributions requiring increased sampling effort to achieve high precision and power (McGarvey et al. 2016, Perkins et al. 2016). Aggregation could be the result of covariate associations and the patchiness of those covariates or to behavioral aspects of species, such as schooling. Our spatial point process modeling approach aims to quantify both these aspects for incorporation into our simulations.

Many of the covariate effects for brown rockfish varied considerably across our study areas making ecological interpretation of these covariates problematic. We found a consistent effect of increasing abundance with depth<sup>2</sup>, indicating a strong depth effect within our study areas. Including surveys at greater depth to establish the limits of this relationship would be informative, and we advise caution in extrapolating these findings outside of

our site boundaries. Other covariates, such as rugosity measures (VRMs and RDMV) and terrain features (BPIs, slope, curvature) have been significant in similar studies for other rockfish species (e.g., Wedding and Yoklavich 2015, Young et al. 2015), although investigations spanning multiple study areas or times are currently lacking. Therefore, whether effects of covariates really differ between study areas due to intrinsic differences or whether effects are largely driven by small sample sizes or temporal variability is currently unclear and requires further research.

Spatial point process modeling indicated that for brown rockfish there was not a strong aggregation of brown rockfish across any of our sites. Where aggregation does occur, sampling effort needs to be higher to achieve the same level of precision in estimates (e.g., McGarvey et al. 2016, Perkins et al. 2016). The spatial parameters found in this study for brown rockfish were similar to other species, including canary rockfish (*Sebastes pinniger*), kelp greenling (*Hexagrammos decagrammus*), and lingcod (*Ophiodon elongatus*) that were modeled across the same sites (*unpublished data*). Practical ranges of spatial correlation for these species were typically in the order of 70–600 m while covariate coefficients were also variable between study areas. Examining the influence of the level of spatial aggregation on sampling precision for these and other species would help inform whether higher sampling effort may be required.

The randomized systematic survey design employed in our study has been shown to be capable of relatively high precision, even when distributions are more highly clustered than those in our simulations (McGarvey et al. 2016). Over many simulations our design covers a wide range of covariates, thereby further reducing the effect of the uncertainty around covariate coefficient importance. Also, the random fields generated in each simulation resulted in different positions for fish at each time step mimicking fish movement between surveys, further reducing the importance of the uncertainty in the coefficients of the covariates for model prediction. The sampling design employed combined with the low level of aggregation for brown rockfish indicates that patterns in power achievable between individual study areas are more likely attributable to differences in starting densities rather than intra-site differences in habitat or spatial aggregation.

#### *Model assumptions and caveats*

Simulation studies inevitably involve assumptions and simplifications during model construction. The expected rate and predictability of the expected recovery in MPAs depend critically on recruitment variability and the ratio of local fishing mortality (local  $F$ ) to the natural mortality ( $M$ ) of the indicator (White et al. 2013). For our chosen indicator species, specific local  $F$  values were not available and thus we used a regional value from stock assessments. Larger or smaller effect sizes due to



differing pre-implementation levels of  $F$  would result in differing times to reach high power to detect change (see Nickols et al. 2019). More refined values for localized  $F$  would also improve predictions made at individual study areas and the ability to test whether recovery rates were meeting these expectations (White et al. 2016, Nickols et al. 2019). We also assumed that natural mortality  $M$  is constant across all sizes, a common assumption in the absence of more detailed data on size-dependent mortality. This assumption would provide a conservative estimate of reserve response and detectability if natural mortality decreases with size as then greater abundances in larger size classes would occur in reserves.

Open population dynamics and constant recruitment were assumed for the projected rate of population change. Brown rockfish have extremely broad larval dispersal (White et al. 2014b), so we made the conservative assumption that there was essentially no local retention of larvae. This assumption is supported by the findings of Nickols et al. (2019) who found that short-term (~10 yr) trends in abundance of a cogenetic rockfish in Californian MPAs were largely driven by interannual variability in recruitment rather than local larval production. Additionally, a genetic parentage analysis of a different cogenetic rockfish species in the Monterey Bay region found evidence for considerable connectivity between MPAs and fished areas within their study domain (Baetscher et al. 2019). Although we did not explore recruitment variability in our simulations, our framework is capable of exploring the effects of different population dynamics through time. In this respect, annual sampling in the early stages of monitoring is likely to be valuable in quantifying recruitment variability and modifying expectations. High interannual recruitment variability is likely to result in smaller population abundance at a given time (see Botsford et al. 2019), which likely will reduce the ability to detect differences between MPA and reference sites.

We kept covariate effects fixed for each study area through time and used posterior medians from our models for the simulations. Taking posterior draws would have allowed integration over the uncertainty in parameter estimates, but this would have also resulted in changes in abundance that were not related to our IPM projections, which were the focus of detection. Where bathymetric covariates can be shown to be consistent predictors, improvements could be made to survey designs in order to better capture populations of a target species across a study area; however, single species are seldom the focus of MPA monitoring programs and improvements in design for one species may come at the cost of another target species.

Spatial models were not utilized in analyzing the time series of simulated data due to the computational demands of such an analysis and the circularity of using the same model for data generation and analysis. These models are likely to provide more precise estimates when applied to empirical data (e.g., Thorson et al. 2015), and

therefore, increased power to detect change. More computationally efficient approaches to spatial modeling are constantly being developed (e.g., see Bachl et al. 2019). Future work examining the utility of these models in a spatiotemporal setting as a time series of data is accrued would be informative.

## CONCLUSION

Our simulation approach provides a powerful tool to explore the interaction between sampling design, spatial distributions of indicators, population dynamics, and metrics for change. While this work does utilize advanced modeling techniques, it also provides an intuitive means of generating simulated data representing recovery inside an MPA. The distributions generated by spatial point processes could be visualized across bathymetric profiles and habitat maps so ecologists could ground truth model outputs for model validation. Also, simulated data could be compared with eventual monitoring data to check if patterns were similar to the empirical data. Spatial point process models are being increasingly applied in ecological settings, and new user-friendly packages are being developed to allow ecologists access to these tools (e.g., see Bachl et al. 2019). The extension of these techniques to other species and survey methodologies can improve and inform design choices for long-term monitoring of MPAs.

## ACKNOWLEDGMENTS

We thank Shauna Oh, Cyndi Dawson, and Becky Ota for implementing a collaboration between U.C. Davis and California Department of Fish and Wildlife and for providing guidance on the utility of this work for management. We thank Katherine Kaplan and Lauren Yamane and the California Department of Fish and Wildlife MPA Management Project team, Stephen Wertz, Adam Frimodig, Michael Prall, Amanda Van Diggelen, Paulo Serpa, Elizabeth Pope, and Sara Worden for conversations and feedback throughout the development of this project. We thank U.C. Davis mentors Louis Botsford and Alan Hastings for feedback and discussion during the course of the project. We thank the team at Marine Applied Research and Exploration (MARE), in particular Dirk Rosen, Andy Lauer-mann and Heidi Lovig for data collection, discussions and feedback during the project. We thank the California Ocean Protection Council and the National Science Foundation (OCE-1909303) for funding this work. Two anonymous reviewers provided helpful comments that improved the manuscript.

## LITERATURE CITED

- Andersen, E. M., R. J. Steidl, and F. Michalski. 2019. Power to detect trends in abundance within a distance sampling framework. *Journal of Applied Ecology* 57:344–353.
- Bachl, F. E., F. Lindgren, D. L. Borchers, J. B. Illian, and R. Freckleton. 2019. inlabru: an R package for Bayesian spatial modelling from ecological survey data. *Methods in Ecology and Evolution* 10:760–766.
- Baddeley, A., E. Rubak, and R. Turner. 2015. *Spatial Point Patterns: Methodology and Applications with R*. Chapman and Hall/CRC Press, London, UK.

- Baetscher, D. S., E. C. Anderson, E. A. Gilbert-Horvath, D. P. Malone, E. T. Saarman, M. H. Carr, and J. C. Garza. 2019. Dispersal of a nearshore marine fish connects marine reserves and adjacent fished areas along an open coast. *Molecular Ecology* 28:1611–1623.
- Banerjee, S., A. E. Gelfand, A. O. Finley, and H. Sang. 2008. Gaussian predictive process models for large spatial data sets. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 70:825–848.
- Barker, D. J., R. H. Clarke, and M. A. McGeoch. 2019. The power to detect regional declines in common bird populations using continental monitoring data. *Ecological Applications* 29:1–15.
- Barrett, N. S., G. J. Edgar, C. D. Buxton, and M. Haddon. 2007. Changes in fish assemblages following 10 years of protection in Tasmanian marine protected areas. *Journal of Experimental Marine Biology and Ecology* 345:141–157.
- Botsford, L. W., J. W. White, M. H. Carr, and J. E. Caselle. 2014. Marine protected area networks in California, USA. *Advances in Marine Biology* 69:205–251.
- Botsford, L. W., J. W. White, and A. Hastings. 2019. *Population dynamics for conservation*. Oxford University Press, Oxford, UK.
- Butler, S. J., R. P. Freckleton, A. R. Renwick, and K. Norris. 2012. An objective, niche-based approach to indicator species selection. *Methods in Ecology and Evolution* 3:317–326.
- Caselle, J. E., and R. B. Cabral. 2018. Monitoring California's rocky marine ecosystems across a network of MPAs: methodological comparison of multiple monitoring techniques. Technical Report to California Ocean Protection Council. Marine Science Institute, University of California, Santa Barbara, California, USA.
- Caselle, J. E., A. Rassweiler, S. L. Hamilton, and R. R. Warner. 2015. Recovery trajectories of kelp forest animals are rapid yet spatially variable across a network of temperate marine protected areas. *Scientific Reports* 5:14102.
- CDFW and OPC. 2018. Marine protected area monitoring action plan. California Department of Fish and Wildlife and California Ocean Protection Council, Belmont, California, USA.
- Claudet, J., D. Pelletier, J. Y. Jouvenel, F. Bachet, and R. Galzin. 2006. Assessing the effects of marine protected area (MPA) on a reef fish assemblage in a northwestern Mediterranean marine reserve: Identifying community-based indicators. *Biological Conservation* 130:349–369.
- Cope, J., E. J. Dick, A. MacCall, M. Monk, B. Soper, and C. Wetzel. 2015. Data-moderate stock assessments for brown, China, copper, sharpchin, striptail, and yellowtail rockfishes and English and rex soles in 2013. National Marine Fisheries Service, Seattle, Washington, USA and Santa Cruz, California, USA.
- Di Stefano, J. 2003. How much power is enough? Against the development of an arbitrary convention for statistical power calculations. *Functional Ecology* 17:707–709.
- Diggle, P. J. 1983. *Statistical analysis of spatial point patterns*. Academic Press, London, UK.
- Easterling, M. R., S. P. Ellner, and P. M. Dixon. 2000. Size-specific sensitivity: Applying a new structured population model. *Ecology* 81:694–708.
- Edgar, G. J. et al 2014. Global conservation outcomes depend on marine protected areas with five key features. *Nature* 506:216–220.
- Ellner, S. P., D. Z. Childs, and M. Rees. 2016. *Data-driven modelling of structured populations: A practical guide to the Integral Projection Model*. Springer, Cham, Switzerland.
- Fairweather, P. G. 1991. Statistical power and design requirements for environmental monitoring. *Australian Journal of Marine and Freshwater Research* 42:555–567.
- Ficetola, G. F., A. Romano, S. Salvidio, and R. Sindaco. 2018. Optimizing monitoring schemes to detect trends in abundance over broad scales. *Animal Conservation* 21:221–231.
- Field, S. A., A. J. Tyre, and H. P. Possingham. 2005. Optimizing allocation of monitoring effort under economic and observational constraints. *Journal of Wildlife Management* 69:473–482.
- Finley, A., H. Sang, S. Banerjee, and A. Gelfand. 2009. Improving the performance of predictive process modeling for large datasets. *Computational Statistics & Data Analysis* 53:2873–2884.
- Foster, S. D., and M. V. Bravington. 2012. A Poisson-Gamma model for analysis of ecological non-negative continuous data. *Environmental and Ecological Statistics* 20:533–552.
- Friedlander, A. M., Y. Golbuu, E. Ballesteros, J. E. Caselle, M. Gouezo, D. Olsudong, and E. Sala. 2017. Size, age, and habitat determine effectiveness of Palau's Marine Protected Areas. *PLoS ONE* 12:e0174787.
- Guidetti, P. 2006. Marine reserves reestablish lost predatory interactions and cause community changes in rocky reefs. *Ecological Applications* 16:963–976.
- Haggarty, D. R., J. B. Shurin, and K. L. Yamanaka. 2016. Assessing population recovery inside British Columbia's Rockfish Conservation Areas with a remotely operated vehicle. *Fisheries Research* 183:165–179.
- Hamilton, S. L., J. E. Caselle, D. P. Malone, and M. H. Carr. 2010. Incorporating biogeography into evaluations of the Channel Islands marine reserve network. *Proceedings of the National Academy of Sciences USA* 107:18272–18277.
- Hatch, S. A. 2003. Statistical power for detecting trends with applications to seabird monitoring. *Biological Conservation* 111:317–329.
- Hayes, K. R. et al 2015. Identifying indicators and essential variables for marine ecosystems. *Ecological Indicators* 57:409–419.
- Hewitt, J. E., and S. F. Thrush. 2007. Effective long-term ecological monitoring using spatially and temporally nested sampling. *Environmental Monitoring and Assessment* 133:295–307.
- Hijmans, R. J., and J. van Etten. 2012. raster: Geographic analysis and modeling with raster data. R package version 2.0-12. <https://rdr.io/cran/raster/>
- Huvene, V. A. I., B. J. Bett, D. G. Masson, T. P. Le Bas, and A. J. Wheeler. 2016. Effectiveness of a deep-sea cold-water coral Marine Protected Area, following eight years of fisheries closure. *Biological Conservation* 200:60–69.
- Kaplan, K. A., L. Yamane, L. W. Botsford, M. L. Baskett, A. Hastings, S. Worden, and J. W. White. 2019. Setting expected timelines of fished population recovery for the adaptive management of a marine protected area network. *Ecological Applications* 29:e01949.
- Karpov, K. A., M. Bergen, and J. J. Geibel. 2012. Monitoring fish in California Channel Islands marine protected areas with a remotely operated vehicle: the first five years. *Marine Ecology Progress Series* 453:159–172.
- Klein, C. J., C. J. Brown, B. S. Halpern, D. B. Segan, J. McGowan, M. Beger, and J. E. M. Watson. 2015. Shortfalls in the global protected area network at representing marine biodiversity. *Scientific Reports* 5:329–337.
- Lecours, V. 2015. Terrain Attribute Selection for Spatial Ecology (TASSE), v. 1.0. [https://www.researchgate.net/publication/314300560\\_TASSE\\_Terrain\\_Attribute\\_Selection\\_for\\_Spatial\\_Ecology\\_Toolbox\\_v\\_1\\_0](https://www.researchgate.net/publication/314300560_TASSE_Terrain_Attribute_Selection_for_Spatial_Ecology_Toolbox_v_1_0)
- Lester, S. E., B. S. Halpern, K. Grorud-Colvert, J. Lubchenco, B. I. Ruttenberg, S. D. Gaines, S. Airame, and R. R. Warner. 2009. Biological effects within no-take marine reserves: a global synthesis. *Marine Ecology Progress Series* 384:33–46.

- Lindenmayer, D. B., and G. E. Likens. 2010. Improving ecological monitoring. *Trends in Ecology & Evolution* 25:200–201.
- Love, M. S., M. M. Yoklavich, and L. Thorsteinson. 2002. *The rockfishes of the northeast Pacific*. University of California Press, Berkeley, California, USA.
- Lyons, J. E., M. C. Runge, H. P. Laskowski, and W. L. Kendall. 2008. Monitoring in the context of structured decision-making and adaptive management. *Journal of Wildlife Management* 72:1683–1692.
- McGarvey, R., P. Burch, and J. M. Matthews. 2016. Precision of systematic and random sampling in clustered populations: habitat patches and aggregating organisms. *Ecological Applications* 26:233–248.
- Moller, J., A. R. Syverseen, and R. P. Waagepetersen. 1998. Log Gaussian Cox processes. *Scandinavian Journal of Statistics* 25:451–482.
- Nichols, J. D., and B. K. Williams. 2006. Monitoring for conservation. *Trends in Ecology & Evolution* 21:668–673.
- Nickols, K. J., J. W. White, D. Malone, M. H. Carr, R. M. Starr, M. L. Baskett, A. Hastings, and L. W. Botsford. 2019. Setting ecological expectations for adaptive management of marine protected areas. *Journal of Applied Ecology* 56:2376–2385.
- Niemi, A., and C. Fernández. 2010. Bayesian spatial point process modeling of line transect data. *Journal of Agricultural, Biological, and Environmental Statistics* 15:327–345.
- Perkins, N. R., S. D. Foster, N. A. Hill, and N. S. Barrett. 2016. Image subsampling and point scoring approaches for large-scale marine benthic monitoring programs. *Estuarine, Coastal and Shelf Science* 176:36–46.
- Perkins, N. R., S. D. Foster, N. A. Hill, M. P. Marzloff, and N. S. Barrett. 2017. Temporal and spatial variability in the cover of deep reef species: Implications for monitoring. *Ecological Indicators* 77:337–347.
- R Core Team. 2019. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. [www.R-project.org](http://www.R-project.org)
- Regan, H. M., M. Colyvan, and M. A. Burgman. 2002. A taxonomy and treatment of uncertainty for ecology and conservation biology. *Ecological Applications* 12:618–628.
- Shanks, A. L., and G. L. Eckert. 2005. Population persistence of California Current fishes and benthic crustaceans: a marine drift paradox. *Ecological Monographs* 75:505–524.
- Skalski, J. R. 2012. Estimating variance components and related parameters when planning long-term monitoring programs. Pages 174–199 in *R. A. Gitzen, J. J. Millspaugh, A. B. Cooper, and D. S. Licht, editors. Design and analysis of long-term ecological monitoring studies*. Cambridge University Press, Cambridge, UK.
- Starr, R. M. et al 2015. Variation in responses of fishes across multiple reserves within a network of marine protected areas in temperate waters. *PLoS ONE* 10:e0118502.
- Strayer, D. L. 1999. Statistical power of presence-absence data to detect population declines. *Conservation Biology* 13:1034–1038.
- Thorson, J. T., A. O. Shelton, E. J. Ward, and H. J. Skaug. 2015. Geostatistical delta-generalized linear mixed models improve precision for estimated abundance indices for West Coast groundfishes. *ICES Journal of Marine Science* 72:1297–1310.
- Tierney, L. 1994. Markov chains for exploring the posterior distributions. *Annals of Statistics* 22:1701–1762.
- Urquhart, N. S. 2012. The role of monitoring design in detecting trend in long-term ecological monitoring studies. Pages 151–173 in *R. A. Gitzen, J. J. Millspaugh, A. B. Cooper, and D. S. Licht, editors. Design and analysis of long-term ecological monitoring studies*. Cambridge University Press, Cambridge, UK.
- Urquhart, N. S., W. S. Overton, and D. S. Birkes. 1993. Comparing sampling designs for monitoring ecological status and trends: Impact of temporal patterns. Pages 71–85. in *V. Barnett, and K. F. Turkman, editors. Statistics for the environment*. Wiley, Lisbon, Portugal.
- USGS. 2019. California State Waters Map Series Data Catalog, Compiled by Nadine E. Golden. [https://pubs.usgs.gov/ds/781/Waagepetersen, R. 2004. Convergence of posteriors for discretized log Gaussian Cox processes. \*Statistics & Probability Letters\* 66:229–235.](https://pubs.usgs.gov/ds/781/Waagepetersen, R. 2004. Convergence of posteriors for discretized log Gaussian Cox processes. Statistics & Probability Letters 66:229–235.)
- Wedding, L., and M. M. Yoklavich. 2015. Habitat-based predictive mapping of rockfish density and biomass off the central California coast. *Marine Ecology Progress Series* 540:235–250.
- Weiser, E. L., J. E. Diffendorfer, L. López-Hoffman, D. Semmens, and W. E. Thogmartin. 2018. Consequences of ignoring spatial variation in population trend when conducting a power analysis. *Ecography* 42:836–844.
- White, J. W. et al 2019. Connectivity, dispersal, and recruitment: connecting benthic communities and the coastal ocean. *Oceanography* 32:50–59.
- White, J. W., L. W. Botsford, A. Hastings, M. L. Baskett, D. M. Kaplan, and L. A. K. Barnett. 2013. Transient responses of fished populations to marine reserve establishment. *Conservation Letters* 6:180–191.
- White, J. W., K. J. Nickols, D. Malone, M. H. Carr, R. M. Starr, F. Cordoleani, M. L. Baskett, A. Hastings, and L. W. Botsford. 2016. Fitting state-space integral projection models to size-structured time series data to estimate unknown parameters. *Ecological Applications* 26:2675–2692.
- White, J. W., A. Rassweiler, J. F. Samhour, A. C. Stier, and C. White. 2014a. Ecologists should not use statistical significance tests to interpret simulation model results. *Oikos* 123:385–388.
- White, J. W., J. Schroeger, P. T. Drake, and C. A. Edwards. 2014b. The value of larval connectivity information in the static optimization of marine reserve design. *Conservation Letters* 7:533–544.
- Young, M., M. H. Carr, and M. Robertson. 2015. Application of species distribution models to explain and predict the distribution, abundance and assemblage structure of nearshore temperate reef fishes. *Diversity and Distributions* 21:1428–1440.
- Yuan, Y., F. Bachl, F. Lindgren, J. B. Illian, S. T. Buckland, H. Rue, and T. Gerrodette. 2016. Point process models for spatio-temporal distance sampling data. *arXiv preprint arXiv:1604.06013*.

## SUPPORTING INFORMATION

Additional supporting information may be found online at: <http://onlinelibrary.wiley.com/doi/10.1002/eap.2215/full>

## DATA AVAILABILITY

Data and example code are publicly available on Zenodo and github: <https://doi.org/10.5281/zenodo.3923229>, <https://github.com/nperkins-utas/MPA-simulation>.