Curtis K Alexandra (Orcid ID: 0000-0001-7284-944X)

[4940]-□
Received: 29 February 2020 | Accepted: 31 August 2020

Running head: CURTIS ET AL.

ARTICLE

Abundance, survival, and annual rate of change of Cuvier's beaked whales (*Ziphius cavirostris*) on a Navy sonar range

K. Alexandra Curtis[1,2] | Erin A. Falcone[3] | Gregory S. Schorr[3] | Jeffrey E. Moore[2] | David J. Moretti[4] | Jay Barlow[2] | Erin Keene[3]

[1]Ocean Associates, Inc., under contract to Southwest Fisheries Science Center, National Marine Fisheries Service, National Oceanic and Atmospheric Administration, La Jolla, California

[2]Marine Mammal and Turtle Division, Southwest Fisheries Science Center, National Marine Fisheries Service, National Oceanic and Atmospheric Administration, La Jolla, California

[3]Marine Ecology & Telemetry Research, Seabeck, Washington

[4]Naval Undersea Warfare Center Division, Newport, Rhode Island

Correspondence

Dr. K. Alexandra Curtis, Marine Mammal and Turtle Division, Southwest Fisheries Science Center, National Marine Fisheries Service, National Oceanic and Atmospheric Administration, 8901

La Jolla Shores Drive, La Jolla, CA 92037.

Email: alex.curtis@noaa.gov

ABSTRACT

Bayesian mark-recapture estimates of survival, abundance, and trend are reported for Cuvier's beaked whales (*Ziphius cavirostris*) using a Navy training range off southern California. The deep-diving beaked whale family is exceptionally vulnerable to mid-frequency active sonar (MFAS), which has been implicated in mass strandings and altered foraging behavior. Extremely low sighting probabilities impede stud of population-level impacts of MFAS on beaked whales. The San Nicolas Basin hosts a Navy training range subject to frequent MFAS use and attracts high densities of *Z. cavirostris*. An 11-year (2007–2018) photo-identification program leveraged automated acoustic detection and location capabilities on the range's 1,800-km^2 hydrophone array to enhance capture probability. Estimated population parameters for *Z. cavirostris* using the range included mean (90% credibility intervals) apparent annual survival of 0.950 (0.899–0.986), annual number of individuals as 121 (71–219), and annual rate of change of −0.8% (−5.6%–4.1%). Simulations show the probability of detecting abundance changes is currently low, but can be greatly improved through continued

monitoring and increased effort. Complementary data collection on habitat use and demographic rates in San Nicolas and surrounding basins is also essential to relating direct effects of MFAS use to changes in vital rates and broader population outcomes.

KEYWORDS

Bayesian mark-recapture, California, Cuvier's beaked whale, sonar, inference error, photo-identification, *Ziphius cavirostris*

1 | INTRODUCTION

Anthropogenic noise in the world's oceans, from ship traffic, seismic surveys, military activity, and other sources, can have negative effects on a broad diversity of marine life (Williams et al., 2015). In cetaceans, which rely on sound to communicate and to sense prey and their environment, acoustic disturbance has been linked to impacts ranging from disruption of communication and feeding behavior to direct mortality (Harwood et al., 2016; Nowacek, Thorne, Johnston, & Tyack, 2007; National Research Council, 2003; Richardson, Greene, Malme, & Thomson, 1995).

Military exercises that use high-powered mid-frequency active sonar (MFAS) for target detection have been associated with mass strandings of beaked whales (Family Ziphiidae) in several areas of the world (Cox et al., 2006; D'Amico et al., 2009; Filadelfo et al., 2009a). Cuvier's beaked whales (*Ziphius cavirostris*) are recorded in mass strandings associated with MFAS use more often than any other beaked whale species (D'Amico et al., 2009; de Quirós et al., 2019). *Z. cavirostris* have the most cosmopolitan distribution of the beaked whales, but they

may also be particularly sensitive to MFAS. Even in a family of prolific divers, these mammalian record-holders for both dive length and depth stand out, routinely conduct foraging dives deeper than 800 m and longer than 60 min (Schorr, Falcone, Moretti, & Andrews, 2014). While the link between diving behavior and strandings corresponding with MFAS use is still unknown, individual or population-level factors may increase susceptibility of some *Z. cavirostris* to decompression sickness (de Quirós et al., 2019).

In the Southern California Bight, intensive MFAS use on U.S. Navy training and testing ranges coincides with localized high densities of beaked whales, including *Z. cavirostris* (Baumann-Pickering et al., 2014; Falcone et al., 2009). Yet no mass strandings have been reported there, nor has a relationship been found for this area between individual strandings and MFAS use (Filadelfo et al., 2009a,b). While this does not rule out the occurrence of MFAS-associated strandings in the region (Faerber & Baird, 2010), MFAS impacts may be mediated by habitat variables, such as bathymetry (D'Amico et al., 2009; Filadelfo et al., 2009a), or by behavioral adaptation (Falcone et al.,

2017).

Beaked whales are notoriously challenging to study due to their offshore distribution, deep and long foraging dives, and cryptic surface behavior. The Southern California Anti-Submarine Warfare Range (SOAR), covering most of the San Nicolas Basin west of San Clemente Island (Figure 1), sees frequent MFAS use and hosts high densities of *Z. cavirostris* (Falcone et al., 2009). After several years of research, including photo-identification work, a dedicated acoustic-visual monitoring program for *Z. cavirostris* at SOAR was established in 2010. The program leverages the Navy's Marine Mammal Monitoring on Navy Ranges (M3R) system to guide sampling effort over the range's 1,800-km^2 array of bottom-mounted hydrophones (Falcone et al., 2009; Morrissey, Ward, DiMarzio, Jarvis, & Moretti, 2006). The resulting augmented encounter rates have enabled focused study of these elusive whales via photo-identification and satellite telemetry.

Dive behavior data from this program have permitted researchers to link MFAS exposure to disruption of *Z. cavirostris* foraging (Falcone et al., 2017), in line with

previously observed behavioral responses to simulated sonar (DeRuiter et al., 2013). These observations join a growing body of evidence for the prevalence of sublethal effects of MFAS on beaked whales (McCarthy et al., 2011; Tyack et al., 2011). Sublethal effects, such as reduced foraging, suppressed reproduction, and reduced calf survival, may dominate the impacts of MFAS on some populations of *Z. cavirostris* (Falcone et al., 2017).

Satellite telemetry and photo-identification results revealed that at least some *Z. cavirostris* individuals return to the San Nicolas Basin repeatedly over weeks, months, and years (Falcone et al., 2009; Falcone & Schorr, 2014), indicating that mark-recapture analysis, and thus local assessment, may be feasible. Local assessment may provide further insight to population-level impacts of observed sublethal effects. However, despite increased capture probabilities made possible by use of the M3R system to localize individual and groups of *Z. cavirostris* on SOAR, capture histories in this data set remain sparse. Variable effort also threatened to confound potential population trend.

Detectability is a key source of variation in the data underlying most population assessments of wildlife (MacKenzie, Nichols, Sutton, Kawanishi, Bailey, 2005). Adequate accounting for variation in detectability is therefore of primary importance to achieving accurate assessment with appropriate estimation uncertainty. In populations with sparse data, decisions on how to aggregate data and which sources of variation to prioritize with limited potential parameters become critical. For mark-recapture data, a number of data aggregation and modeling approaches may aid in inference of population parameters from sparse capture histories. Several common approaches constitute a form of information-sharing across strata (MacKenzie et al., 2005), including: aggregating information from multiple occasions per year to annual-level capture histories, representing individual or temporal heterogeneity in capture probability as a parametric distribution in a hierarchical model (e.g., Dorazio & Royle, 2003), and using a covariate to parameterize heterogeneity in capture probability. In distance sampling, detection probability is commonly modeled as a function of sea state, a concept that

was recently extended to modeling detection probability on the trackline, of particular importance for less conspicuous species (Barlow, 2015).

We present estimates of annual apparent survival, recent annual abundance, and annual rate of change of *Z. cavirostris* using (i.e., inhabiting or passing through) SOAR, based on eleven years of photo-identification data. We used a Bayesian framework to estimate parameters for hierarchical models with an effort covariate and random effects for capture probability, allowing us to explain a key source of variability and capture important sources of uncertainty. We used a recent methodological advance that integrates data from both left- and right-side captures (or, more generally, from multiple noninvasive marks; Bonner & Holmberg, 2013; McClintock, 2013) through latent multinomial capture histories (Link, Yoshizaki, Bailey, & Pollock, 2010), thereby augmenting capture probability without inflating precision. Motivated by the need to preserve reliable individual identification, requiring a high-quality photo for each individual, without omitting medium-quality recaptures, we developed a novel data selection approach to

accommodate different capture and recapture probabilities (i.e., trap response) in trend estimation. Finally, we used simulation to inform recommendations for continued monitoring of *Z. cavirostris* in SOAR.

2 | METHODS

2.1 | Field work and photo identification

Photo-identification data for *Z. cavirostris* were collected at SOAR in the San Nicolas Basin (Figure 1) from 2006 to 2018 (Falcone & Schorr, 2014; Falcone et al., 2009;). Data collection involved close collaboration between a field team, operating in one or two 5- to 7-m rigid-hull inflatable boats, and the M3R passive-acoustic observer team, who monitored the 1,800-km^2 hydrophone array on SOAR remotely from shore. The M3R team interpreted automated acoustic detections and locations of cetacean vocalizations by the M3R system and provided the information to the field team (Jarvis, Morrissey, Moretti, DiMarzio, & Shaffer, 2014; Moretti, 2015). Each time *Z. cavirostris* were encountered, with or without M3R direction, the field team collected photo-identification and other data. During effort, the GPS location of the vessel, depth, Beaufort sea

state, visibility, and swell height were recorded continuously. Over the course of the monitoring program, field effort occurred in every month except September and December. Effort since 2010 focused on winter and spring months to complement the 2006–2009 effort, which occurred predominantly in the late summer and fall, and to maximize encounter rates, since density of *Z. cavirostris* on SOAR appears to decline in late summer and early fall (Baumann-Pickering, Hildebrand, Yack, & Moore, 2015).

Photographs from each *Z. cavirostris* encounter were reviewed to determine the number of unique individuals present. At least one image including the dorsal fin, and up to four total images (anterior and dorsal region views of each side of the body, as available), of each individual were selected for complete photo-identification processing. Each of these images were scored for four aspects of quality (proportion of the body visible, angle, sharpness, and exposure) using a three-point scale (1 = best to 3 = worst), which were then averaged to an overall photo quality score. Individual animals were scored for distinctiveness each time they were photographed based on the number of distinctive markings (e.g., pigmentation patterns or

scars) visible in the best photograph of each side of the body, with scores ranging from 1 (no marks or scars) to 5 (20 or more marks). Whales were assigned to age and sex classes using a combination of behavioral (e.g., mother-calf associations), genetic (for individuals that were biopsy sampled during the study), and appearance data (e.g., erupted teeth, scarring rates) following methods described in Rosso, Balardini, Moulins, and Würtz (2011) and Coomber, Moulins, Tepsich, and Rosso (2016). Individual sighting histories were created through the comparison of captured individuals in each sighting to the accrued catalog, as described in Falcone and Schorr (2014).

2.2 | Data selection

We limited the scope of this study to sightings within the San Nicolas Basin (Figure 1), and to photo-identifications from October 2007 onwards, where after the vectoring guidance from the acoustics team improved in accuracy, and the field team was able to reach the full SOAR area for sampling, including favorable habitat for *Z. cavirostris* that had previously been undersampled (Falcone & Schorr 2014; Falcone et al., 2009).

Two photo quality criteria and one distinctiveness

criterion were used in building the mark-recapture data set (Figure 2): (1) a higher photo quality criterion ("catalog-qualifying") for inclusion of an individual in the data set, (2) a slightly lower bar for photo quality ("recapture-qualifying") for inclusion of further encounters with that individual in its capture history, and (3) a distinctiveness criterion ensuring reliable identification in recaptures. A photograph was catalog-qualifying if it scored ≤2 for all four photo quality aspects. This threshold enabled reliable identification even for individuals lacking distinctive marks or scars (based on variation in fin shape, pigmentation, and other subtle characteristics). Therefore, since the field team did not discriminate among whales during photographic sampling, all encountered whales were equally likely to be included as identified individuals in the catalog regardless of age and sex, with the caveat that the capture probability of calves may be reduced by their small size and/or position relative to the mother. Recapture-qualifying photographs were those with a mean photo quality score ≤2, which could reliably be matched to the catalog-qualifying photo for distinctiveness scores ≥2, so these

thresholds of quality and distinctiveness were used to further filter records from the catalog to create a relatively homogeneous population in terms of identifiability. The difference between recapture-qualifying and catalog-qualifying thresholds led to a difference between initial capture and recapture probabilities, equivalent to a trap response. The implications for each mark-recapture model framework are discussed in their respective sections below. This approach allowed us to maximize our sample size without introducing unquantifiable sources of bias due to variability in distinctiveness, which changes with age and sex.

Records of calves were excluded from mark-recapture analysis, because the association between a calf and its mother violates the mark-recapture assumption of independent fates, and because we expected survival rates of calves to be the most different from other age classes.

Due to the relatively small size of the photo-ID data base, we were able to screen records retrospectively for matches after a catalog-qualifying record was obtained for an individual. We termed these recapture-qualifying records that preceded the

catalog-qualifying record “precaptures” (elsewhere referred to as retrospective recaptures; Pace, Corkeron, & Kraus, 2017). We created two data sets for mark-recapture analysis: one which included precaptures and one which omitted them. Use of each data set in estimating population parameters (Figure 2) was guided by simulations (Appendix S1) and sensitivity tests, as detailed below.

Capture data were aggregated into eleven annual occasions from 2007 to 2018, with each annual occasion beginning August 1, which corresponded to a natural break in timing of sightings across years and fell relatively close to the start date for data included in the analysis (October 2007).

2.3 | Sex and age composition

Quality and distinctiveness thresholds were relaxed as follows to calculate population composition. Sex and age class composition were calculated as mean annual proportions of individual annual captures, from recapture-qualifying sightings of all individuals with catalog-qualifying photos (i.e., not filtered for distinctiveness or age class). Animals can be classified as calves or not with confidence in recapture-

qualifying photos, so we estimated reproductive rate as the proportion of all otherwise unfiltered recapture-qualifying sightings that were staged as calves (Steiger & Calambokidis, 2000), with Wilson-score binomial confidence intervals.

2.3 | Model development

A Jolly-Seber model, which would have estimated all three population parameters of interest, failed to converge, and the associated assumption of equal capture and recapture probabilities was also violated. We used three separate Bayesian mark-recapture models to estimate apparent annual survival, annual abundance, and annual rate of change. All three models jointly estimate capture probability and the respective parameter of interest. Our mark-recapture model development focused on accounting for individual and temporal heterogeneity in capture probability p (e.g., MacKenzie et al., 2005). We first describe model development for this shared parameter before providing further detail on the three specific models used to estimate the parameters of interest.

Previous work on *Z. cavirostris* found age and sex effects on extent of marking (e.g., Rosso et al., 2011) and sex effects

on inter-sighting intervals (McSweeney, Baird, & Mahaffy, 2007). To assess sex-associated heterogeneity in capture probability, survival, and site fidelity, we calculated sex-specific means of individual (1) number of annual captures, (2) number of sightings, (3) mean interval between annual captures for recaptured individuals, and (4) length of positive capture history. Mean statistics for males and females were well within one standard error of each other (in most cases almost identical), so sex was not considered further in model development.

Satellite telemetry of *Z. cavirostris* in the San Nicolas Basin suggests individuals vary in the proportion of time spent in SOAR, and thus in their availability for encounter (Schorr et al., 2014). To allow for individual heterogeneity, which can otherwise lead to severe underestimates of population size (Carothers, 1973; Cormack, 1972), we considered a term for individual random effects in capture probability.

Social groups and accompanying nonindependence of detection may further inflate variance in capture probability and lead to biased abundance estimates, even when individual heterogeneity

is included in the estimator or model (Boulanger, McLellan, Woods, Proctor, & Strobeck, 2004; Gupta, Joshi, & Vidya, 2017). Little is known about the social structure of beaked whales, but some evidence exists that association in *Z. cavirostris* is fluid (McSweeney et al., 2007). Repeat associations among individuals over multiple years account for only a small fraction of sightings of *Z. cavirostris* at SOAR, half of which were mother-calf pairs (see Results). Calves were omitted from mark-recapture analysis (see Data selection), and the few other observed cases of association were of short duration compared to the time series. Association was not considered further in model development.

Effort varied considerably over the time series in terms of total hours and seasonal allocation of sampling (Figure 3), including the aforementioned shift from summer and early fall months, when density of *Z. cavirostris* in SOAR appears to decline (Baumann-Pickering et al., 2015; D.J.M., unpublished data) relative to other months. We considered two different effort indices as temporal covariates to account for the expected temporal heterogeneity in capture probability. For both

metrics, we only included time spent searching within the San Nicolas Basin over bottom depths >800 m, since no sightings occurred at shallower depths (Figure 1).

The vast majority of sightings occurred in relatively calm conditions, so the first effort metric (β_1) was simply annual hours of search effort in Beaufort Sea States 0 through 2. The second effort metric (β_2), annual expected captures, was designed to account for the strong expected and observed pattern in capture rates (r) with Beaufort sea state and season (Figure 4):

$$r_{b,s} = \frac{\sum_y c_{y,b,s}}{\sum_y h_{y,b,s}}$$

where c is captures, defined as all recapture-qualifying records, regardless of inclusion in the data set for analysis (i.e., not filtered for individuals with catalog-qualifying records, by age class, or by distinctiveness), h is effort hours, y is year, s is season, and b is Beaufort sea state. We calculated expected annual captures as the summed product of the rate for each b-s bin and the corresponding annual effort:

$$\beta_{2,y} = \sum_{b,s} h_{y,b,s} r_{b,s}$$

Both effort indices were standardized to a mean of zero and standard deviation of one. We also included a term for temporal (annual) random effects in capture probability where model flexibility allowed (i.e., for estimating population growth rate, as it is not an available term in the *multimark* models fit for survival and abundance), to account for remaining unexplained temporal variability in capture probability.

To guide model choice with respect to parameterization of capture probability, we used reversible-jump Markov chain Monte Carlo (RJMCMC) sampling of candidate model fits of open-population Cormack-Jolly-Seber (CJS) models in the *multimark* package in R (Cormack, 1964; Jolly, 1965; McClintock, 2015; R Core Team, 2016; Seber, 1965). The *multimark* package implements mark-recapture models that accommodate capture histories for two non-invasive marks simultaneously, in a Bayesian framework. We fit *multimark* CJS models to left- and right-side capture histories for the 11-year time series, followed by RJMCMC sampling of the resulting fits of candidate models. The CJS

model selection results informed model selection for the other two mark-recapture frameworks (for estimation of abundance and rate of change), which each used a smaller subset of the data (see below and Figure 2).

The CJS model framework, which estimates capture probability (p) and apparent survival (ϕ), conditions on first capture and is unaffected by a difference between initial capture and recapture probabilities. Including precaptures may introduce a small positive bias in ϕ (<0.01; see Appendix S1), but preliminary analysis showed little or negative difference in estimates of ϕ for the data set including precaptures compared to that excluding them. We therefore included precaptures in the data set used for model selection. We considered four models of increasing complexity for capture probability (p): an intercept-only model, intercept plus individual random effects, and intercept plus individual random effects plus one or the other effort covariate. Apparent annual survival, which includes death and emigration, was modeled as constant among years and individuals. We assumed an equal probability of photographing either side of an individual in an encounter. The models in the

multimark package use known links between left and right-side capture histories to estimate a latent multinomial. We classified capture histories for left and right sides as known to be linked for individuals with filtered photo-IDs from both sides within a sighting or characterized by distinctly shaped or damaged dorsal fins that allowed reliable bilateral identification. Additional details on model, priors, and MCMC and RJMCMC sampling are in Appendix S2.

2.5 | Estimation of survival

We used the data set including precaptures and the *multimark* CJS model identified as the most probable through RJMCMC sampling, as described and justified in the previous section, to estimate annual apparent survival ϕ. Further details are in Appendix S2.

2.6 | Estimation of abundance

To estimate the annual number of *Z. cavirostris* using SOAR, we fitted a closed population model in *multimark* to the three most recent years of annual-level capture histories, August 2015 through July 2018 (Huggins, 1989; McClintock, 2015). Capture probability was parameterized as for the CJS model chosen through RJMCMC as described above. Simulation showed that

including precaptures when fitting closed-population models could lead to substantial underestimation of abundance (see Appendix S1), because it interferes with correct estimation of initial and recapture probabilities, and capture probability is the linchpin of mark-recapture abundance estimation. We excluded precaptures for this analysis. Closed-population models allow separate estimation of initial and recapture probabilities, but this parameterization resulted in a slightly higher estimate for initial than recapture probability, so we used a single-intercept model for capture probability. Further details are in Appendix S2.

We used the approach of Wilson, Hammond, and Thompson (1999) to estimate a correction factor to calculate noncalf and total abundances from the model estimate, which was based on a filtered subset of animals. We applied their method at the annualized capture record level, to all individuals with catalog-qualifying records (regardless of age class or distinctiveness). We corrected for annual mean proportions of (1) noncalf animals that were insufficiently marked to be included in the analysis, to obtain a noncalf population

estimate (N_{nc}); and (2) calves and insufficiently marked animals, for a total population estimate (N).

2.7 | Estimation of annual rate of change

To estimate the mean annual rate of change in abundance of *Z. cavirostris* using SOAR, we fit single-side capture histories with a Bayesian Pradel-lambda model, which estimates λ (annual rate of change plus one, henceforth "population growth rate"), ϕ, and p (Pradel, 1996; Tenan et al., 2014). The Pradel-lambda mark-recapture model framework uses reversed capture histories to estimate population growth rate in the same manner that ϕ is estimated from forward-moving capture histories in a CJS model. Given greater recapture than initial capture probability, λ will tend to be overestimated by an amount that depends on the magnitude of that difference (Hines & Nichols, 2002). Simulations for our parameter space showed that λ may be positively biased by as much as 0.01 to 0.02, but that this bias is eliminated when precaptures are included in the data (see Appendix S1), so we included precaptures for this analysis. Capture probability was modeled as above, but with temporal (annual) rather than individual random effects. Individual

random effects do not bias estimation of constant population growth rate in mark-recapture models (Hines & Nichols, 2002). To incorporate the better-informed estimate of ϕ from the *multimark* CJS model in our estimate of λ, we fixed ϕ as a constant in the Pradel-lambda model, and used multiple imputation to estimate the other model parameters over successive values of ϕ sampled from the posterior distribution of ϕ from the CJS model. We adapted the BUGS code written by Tenan and colleagues (2014) by (1) fixing the value of ϕ to a constant; (2) moving the temporal random effects term from λ to p; (3) moving the temporal covariate from ϕ to p; and (4) changing the prior for the logit-transformed intercept for p to (0, 10), because our parameter space is near the boundaries and could thus be unduly influenced by the original prior of (0, 100). See Appendix S2 for more details.

We fitted the model with the *rjags* package in R (Plummer, 2016, R Core Team, 2016). We used right-side capture histories, because in preliminary analyses with a Bayesian CJS model, posteriors for apparent survival were better resolved from the upper boundary of one for the right-side capture histories than

for the left. We used data cloning with CJS models in MARK through RMark to ensure that the parameters near boundaries (ϕ and p) are identifiable (Cooch & White, 2016; Lele, Dennis, & Lutscher, 2007; Lele, Nadeem, & Schmuland, 2010). For each of 250 sampled values of ϕ, we fitted a Pradel-lambda model, and the resulting posteriors for λ were combined to obtain an aggregate result. Further details are in Appendix S2. To evaluate the effect of including the effort covariate, we fitted the same Pradel-lambda model (with ϕ as an estimated parameter with the same prior as p), with a vector of zeros instead of the annual effort index.

2.8 | Convergence diagnostics

All model fits were evaluated for mixing and convergence by examining trace plots, running mean plots, and potential scale reduction factors (psrf) for each parameter with the *coda* and *mcmcplots* packages in R (Curtis, 2015; Plummer, Best, Cowles, & Vines, 2006). In all cases, trace plots showed good mixing of all three chains, running mean plots demonstrated stable results, and psrf was $\leq$1.01 for both point estimates and credible intervals.

For the multiple-imputation fit of the Pradel-lambda model, we examined running mean plots and successive density plots of the posterior distribution for λ with increasing numbers of imputations to assess sufficiency of the number of imputations used.

2.9 | Inference error and sampling design

We used simulation to provide insights into our current and potential future ability to detect a decline in abundance. Understanding potential errors in inference is critical in a decision-making framework, where management action or lack thereof may result in overprotective (inferring a decline when there is none) or underprotective (failing to infer a true decline) management errors, respectively. We thus refer to errors in inference as over- or underprotective. We focused on inference from the Pradel-lambda model described above, with an effort covariate and temporal random effects for capture probability, but with survival as one of the parameters to be estimated.

We considered 14 scenarios and simulated 1,000 sets of capture histories for the population for each scenario.

Simulated populations experienced one of two constant rates of decline, corresponding to 50% declines over 20 or 10 years (i.e., 21 or 11 years of mark-recapture), or λ of 0.966 (ϕ = 0.95) and 0.933 (ϕ = 0.91), respectively. Other parameters were based on posteriors from the preceding analyses. The first two scenarios served as "base case" scenarios, one each for the two rates of decline, with 11 years of simulated sampling at current effort/capture probability levels, including a known effort covariate as well as temporal and individual random effects, all three of which were modeled as draws from normal distributions with means of zero and specified variances. Scenarios 3 through 12 varied duration of simulated sampling and annual effort. In Scenario 13, we explored our ability to detect an extreme rate of decline in abundance of 75% over the 11-year capture history (λ = 0.87, ϕ = 0.85). In Scenario 14, we set λ = 1 to characterize probability of overprotective error for our current data complement.

Pradel-lambda models were fitted to the resulting simulated data sets for each scenario using three MCMC chains, each including an adaptation phase of 500 samples, burn-in of 1,000

samples, and 10,000 iterations, resulting in 30,000 samples. For each simulation, we tracked whether (1) the model had converged (i.e., psrf < 1.05 for all parameters), (2) percentage of posterior probability distribution of $\lambda < 1$, (3) percentage of time that true φ fell within the 80% and 95% posterior credible intervals, and (4) percentage of time that true λ fell within the 80% and 95% posterior credible intervals. Probability of detecting declines and under- and overprotective error rates were evaluated based on those simulations for which the model converged.

3 | RESULTS

3.1 | Photo identification data

For individuals with catalog-qualifying records (see Methods), mean sightings per individual showed no consistent trend with distinctiveness. The annual-level discovery curves for right-side and left-side filtered capture histories show slowing over the time series, but continue to rise steadily (Figure 5). Capture frequencies for all individuals and encounters included in the analyses are summarized in Table 1. Although most individuals were seen only within one annual occasion, positive

capture histories spanning more than one annual occasion included every possible duration from 2 to 11 years, with 25 individuals having positive capture histories of 5 years or longer. The mean interval between positive occasions for individuals that were recaptured, weighted by individual, was 3.1 years. Individuals with repeated photo-IDs across years showed no seasonal pattern in occurrence. The most recent three years of mark-recapture data from both sides, to which the closed-population *multimark* model was fitted, include 46 capture histories, 37 of which are known for both sides, with a mean of 1.26 captures each.

We examined observations of repeated association among all individuals with catalog-qualifying records, regardless of age class or distinctiveness. Repeated associations among individuals were relatively rare. Repeated associations among three or more individuals (three instances) were only observed within a few days of each other. Repeated associations among individuals over periods longer than 6 months included three between an adult female and a calf, two of which matured to juveniles while in association, which lasted as long as 4.5

years; two between an adult female and an adult male (observed over a maximum of 1.2 years); and one between a subadult male and an adult male (2.4 years).

Group sizes of sighted whales averaged 3.0 (± 1.8 *SD*, 0.2 *SE*), somewhat lower than the mean of 3.8 first reported in Falcone et al. (2009) for the same study area and data collection methods.

3.2 | Sex and age composition

The following data summaries are based on a mean of 18.9 individual captures per occasion over 11 annual occasions. The mean annual sex ratio of annual captures of individuals with catalog-qualifying records is similar to those reported in Falcone et al. (2009), with 45.3% female, 38.5% male, and 16.2% unknown. The mean annual age class distribution of annual captures of individuals with catalog-qualifying records is 61.5% adults, 21.5% subadults, 8.5% juveniles, 4.7% calves, and 3.9% unknown. The reproductive rate, i.e., the proportion of calves among all otherwise unfiltered recapture-qualifying records, was 6.0% (16 of 267 sightings; 95% CI 3.7%, 9.5%). The percentage of calves may exceed annual calf production, since one individual

was staged as a calf in two successive years. On the other hand, these estimates may be biased low, because calves may be less likely to be successfully photographed than older animals due to their smaller size.

3.3 | Model selection

Relative probabilities of the four candidate *multimark* CJS models included in RJMCMC sampling are in Table 2. The most likely model included individual heterogeneity and the effort covariate that accounts for Beaufort Sea State and season. Effective sample size for the model selection parameter was 721, and the multivariate psrf was 1.006.

3.4 | Abundance, trend, and demographic rates

Estimates for survival, slope of capture probability with effort index, abundance, and annual rate of change for non-calf *Z. cavirostris* using SOAR are summarized in Figure 6 and Table 3 (see Appendix S2 for all estimated parameters). For skewed posterior distributions, such as annual apparent survival and abundance, the posterior mode is most analogous to a frequentist maximum likelihood estimate. The abundance estimate has been multiplied by a correction factor that accounts for the

proportion of animals that were calves or did not meet distinctiveness criteria (uncertainty is negligible relative to that of the abundance estimate; 1.052 for the noncalf abundance, 1.107 for total abundance, including calves).

The slope of capture probability with effort has a 99.3% posterior probability of being greater than zero, supporting a positive relationship between effort and capture probability (Figure 6). The closed-population estimate of noncalf abundance for effort years 2015–2017 has a mode of 84 (90% credible interval: 67–208; Table 3), placing it at the low end of earlier adult-only estimates of abundance based on a Lincoln-Peterson estimator and single-side capture histories, which ranged from 101 to 332 with *CV*s of 15% to 29% for several combinations of 2-year intervals from 2007 through 2014 (Falcone & Schorr, 2014). The estimate for λ suggests the abundance is stable or decreasing slowly (mean of 0.8% annual decrease), with 62% of the posterior distribution <1. By comparison, the model with no effort covariate resulted in a mean λ of 1.02 with a 31% posterior probability of being <1, despite also producing a lower estimate of ϕ than the CJS model from which the multiple

imputation samples were drawn, underscoring the importance of accounting for nonrandom sources of variation in capture probability when estimating λ.

3.5 | Inference error and sampling design

Several parameters for simulated capture histories were drawn from posteriors from the above mark-recapture analyses, including *p* (0.085 for current effort, increasing by 40% at doubled effort), starting abundance (100), temporal fixed (i.e., effort) and random effects (σ_{tFE}^2 = 0.2, σ_{tRE}^2 = 0.15), and individual random effects (σ_{iRE}^2 = 0.25). At most four simulations failed to converge (psrf for all parameters <1.05) for any given scenario, and were filtered out before further summarizing the results.

Over- and underprotective inference error rates for data scenarios representative of the current sampling program (Figure 7) show we currently have low probability (high underprotective error rate) of detecting a decline corresponding to a 29% cumulative decrease in abundance (λ = 0.966) over the 11-year time series (10 intervals). A 50% cumulative decrease (λ = 0.933) would not be reliably detected either. There is high

probability of detecting an extreme decline (λ = 0.87), corresponding to a 75% decrease in abundance over the time series.

Simulations with increased effort or additional years of effort (Figure 8) show that doubling effort would have led to modestly increased probability of detecting rates of decline of −3.4% or −6.7% per annum (λ of 0.966 or 0.933) at 10 or 15 years. But at 20 years, when a −3.4% per annum decline (λ = 0.966) would correspond to a 50% cumulative decrease in abundance, all scenarios have fairly high probability of inferring a decline.

4 | DISCUSSION

We analyzed 11 years of photo-ID data from the San Nicolas Basin to produce the first estimates of apparent survival, abundance, and annual rate of change of *Z. cavirostris* on a Navy training range subject to frequent MFAS use. Beaked whales are notoriously challenging to assess due to low sighting probability; for example, Cañadas et al. (2018) had to combine 27 years of visual survey data in the Mediterranean to obtain a single abundance estimate for *Z. cavirostris*. In the California

Current Ecosystem (CCE), Bayesian hierarchical models were used to infer broad-scale abundances and trends for several species of beaked whales from visual transect data (Moore & Barlow, 2013, 2017), but coefficients of variation remained high, circa 70% for *Z. cavirostris*. As such, a more localized assessment, e.g., to investigate impacts of MFAS use within the Southern California Bight, was not possible from their dataset. The results of our mark-recapture and simulation analyses provide robust estimates of local abundance, apparent annual survival, and change in number of *Z. cavirostris* using SOAR, but they also underscore that we do not yet have the sample size to reliably detect even a 50% decline in abundance over the course of the 11-year monitoring program ($\lambda = 0.933$).

The high density of *Z. cavirostris* in SOAR and long-term site fidelity of individuals despite repeated MFAS use suggest that the San Nicolas Basin contains important foraging habitat. Evidence for widespread, long-term site fidelity to the San Nicolas Basin, based on telemetry data (Schorr et al., 2014) and photographic resightings over 5 years or longer for 25 individuals, reinforces the conceptual model of small resident

populations as an important component of *Z. cavirostris* demographics (Allen, Brownell, Yamada, & Mead, 2012). Similar behavior has been observed in beaked whales at other long-term study sites off Hawaii and the Bahamas (Claridge, 2013; McSweeney et al., 2007). The additional support presented here for a resident population in the San Nicolas Basin and immediate vicinity bolsters the rationale for long-term monitoring and assessment of *Z. cavirostris* at SOAR.

Several lines of evidence also support the prevalence of transients passing through SOAR. In the filtered data, we found that individuals only seen on one occasion tended to have a lower probability of multiple sightings per annual occasion than individuals seen over multiple years (0.12 versus 0.16). The discovery curves, though showing initial slowing, continued to rise steadily in the last few years (Figure 5). A CJS model with separate survival rates for the first year after initial capture and subsequent years, which would eliminate negative bias in the survival estimate due to transients that arrive and depart within an annual occasion (Pradel, Hines, Lebreton, & Nichols, 1997), was supported in RJMCMC model selection. However, this

model did not result in estimates of long-term survival that were resolved from one, so it was not presented here. Dedicated simulation work is merited to better understand how transience factors into population assessment for *Z. cavirostris* in SOAR.

The number of *Z. cavirostris* that use SOAR annually is on the order of a few percent of the estimated broader population in the CCE (Moore & Barlow, 2017). The short-term, closed-population abundance estimate that we presented minimizes the influence of transients, but may all the same overestimate the resident population and underestimate the superpopulation for that period of time. Additional potential sources of bias are also towards overestimation: false negatives are usually more likely than false positives in photo-ID studies (Stevick, Palsbøll, Smith, Bravington, & Hammond, 2001), and the use of 3 years of data (2015–2018), rather than one, would permit some turnover due to apparent survival and new recruitment. However, our estimates of survival rate and annual rate of change suggest the latter bias is likely less than 10%.

The estimate of annual survival rate (ϕ ; M = 0.95) is within the range expected for long-lived cetaceans, particularly

considering that the analysis included juveniles that had distinctive marks (8% of filtered annual-level photo-IDs) and the potential for negative bias due to transience. This estimate exceeds that for pooled adult and subadult *Z. cavirostris* off El Hierro, in the Canary Islands of the Northeast Atlantic, where sonar use has been banned within 92.6 km since 2004, which was estimated at 0.91 (95% CI 0.82, 0.95; Suárez, 2018). In Blainville's beaked whales (*Mesoplodon densirostris*) at Abaco, The Bahamas, where sonar use is limited, estimated ϕ is only slightly higher than for *Z. cavirostris* in SOAR, and is heterogeneous among sexes and age classes, with higher ϕ in adults and females than subadults and males (Claridge, 2013). Additional preliminary analysis with the *multimark* package using RJMCMC for model selection supported an intercept-only model for φ, so we were unable to resolve sex or age class differences.

The proximity of the point estimate of λ (annual rate of change plus one) to one, for a time series beginning in 2006, is concordant with the leveled-off population growth rate for the broader population of *Z. cavirostris* in the CCE reported by Moore and Barlow (2017) for the late 1990s to 2014, after a

decline in the early 1990s. Passive acoustic estimates of instantaneous numbers of *Z. cavirostris* in SOAR from the hydrophone array also did not show any trend over 5 years (DiMarzio, Jones, Moretti, Yhomas, & Oedekoven, 2018), but this metric of use intensity of SOAR may not necessarily covary with the overall abundance of animals using the range, as estimated by our mark-recapture analysis. For example, use intensity could exhibit hyperstability (sensu Hilborn & Walters, 1992), remaining constant while the overall abundance decreases, due to the high value of the foraging habitat. The high estimated survival rate noted above is paired with an estimated percentage of calves in the population of 4.7% or 6% depending on approach (with the weight of potential bias towards underestimation, as described in Results), within the range expected for long-lived cetaceans. More information on stage- and sex-specific survival rates and reproductive rates are needed to reliably infer a population trajectory based on vital rates. The direct estimate of λ from existing photo-ID data in SOAR has a 90% credible interval spanning fast decline (0.944) to rapid increase (1.041).

Characterizing the limitations of inference is essential to drawing robust conclusions about population status and trends (Taylor & Gerrodette, 1993). Based on simulation of current and potential future sampling design, we can be fairly certain that *Z. cavirostris* using SOAR are not experiencing a decline in abundance on the order of 13% per year (75% decrease over the time series; Figure 7 and Table 3). But several additional years of effort would be required to reliably detect a 6.7% annual rate of decline, and approximately 10 more years—corresponding to a 50% cumulative abundance decrease—would be required to reliably detect a 3.6% annual rate of decline (Figure 8). Taylor et al. (2007) predicted a Type II error rate of 90% for detecting a 50% decrease in a beaked whale stock from 15 years of standard visual surveys, so the innovative mark-recapture program for *Z. cavirostris* at SOAR has achieved remarkable probability of detecting declines. Nonetheless, greater strength of inference is needed to detect a potential decline before the population has been diminished to such a severe extent.

Inference from photo-ID data may be improved through a number of strategies. Most obviously, capture probability and

thus precision can be increased by increasing annual effort, as evidenced by the conclusively positive relationship between effort and capture probability found in this analysis. Further analytical innovation may also afford greater estimation precision. Extending latent capture history models that integrate capture histories from left and right sides to Pradel-lambda and Jolly-Seber frameworks would effectively increase capture probability and thereby estimation precision, as well as integrate estimation of two or more key parameters from all available photo-ID data within a single model. Combining photo-ID data with other data types by fitting integrated models may also prove to be a valuable approach. For example, capture probability estimation in a Pradel-lambda model may be refined by constraining it with passive acoustics data on year-to-year or seasonal changes in density of *Z. cavirostris* in SOAR.

Collection and integration of different data types may provide far more insight than just increased precision. Even with a more precise estimate of λ, we would not be able to rule out the possibility that SOAR might be acting as a sink (or a source) for the larger population (e.g., Whitehead & Gero,

2015). *Z. cavirostris* in SOAR respond to MFAS with changes in foraging behavior (Falcone et al., 2017). Physiological modeling indicates that sublethal effects of MFAS on beaked whales are most likely to be reflected in reduced fetus and calf survival (New, Moretti, Hooker, Costa, Simmons, 2013). Habituated animals might have reduced or unsuccessful reproduction or recruitment to the reproductive class due to diminished foraging efficiency, or require an increased threshold of prey density to support successful reproduction, corresponding to a diminished carrying capacity. And although strandings have not been observed in association with sonar events off California (Falcone et al., 2017; Filadelfo et al., 2009a,b), direct mortality due to exposure to sonar cannot be ruled out at SOAR, particularly considering the offshore location (Faerber & Baird, 2010). As a result of any of these mechanisms, animals using SOAR might be experiencing gradual replacement from the broader population rather than maintenance through local production, or failing to fulfill a previous role as a source.

Complementary data are critical to characterizing population processes and assessing whether MFAS use is affecting

the vital rates of individuals that use SOAR. Key parameters for which more information is needed include age of maturation, pregnancy rates, calf production, and calf survival. Collecting similar data at a comparable site in the CCE that is minimally exposed to MFAS would provide a valuable comparison (National Academy of Sciences, Engineering, and Medicine, 2017), similar to the role of Abaco in evaluating the impacts of MFAS on *M. densirostris* at AUTEC (Claridge, 2013). Alternatively, collecting joint data for individuals on pregnancy and stress hormones from biopsy, calf accompaniment rates from sightings, and telemetry or, perhaps more tractably, stable isotope profiles (Fleming, Kellar, Allen, & Kurle, 2018) could provide insight into how individual heterogeneity in habitat use—and thus exposure—relates to reproductive success (Gimenez et al., 2018; Pirotta et al., 2018). Improved data on vital rates would support matrix population modeling to elucidate whether local productivity would be sufficient to match observed λ or would require immigration (e.g., Whitehead & Gero, 2015). It may also be possible to increase capture probabilities sufficiently to gain more insight into immigration and emigration through mark-

recapture analysis, but this may be a reach for these elusive animals. Meanwhile, continued monitoring of sex ratio in *Z. cavirostris* at SOAR is valuable not only to inform matrix modeling, but potentially as an independent indicator: changes in sex ratio may be apparent before a clear picture emerges of a downward population trajectory (Pace et al., 2017).

In making the best of available data and methods to obtain population parameter estimates for *Z. cavirostris* using SOAR, we faced a common potential source of bias in population growth rate estimation from photo-ID data. Higher standards for inclusion of an individual in a catalog than for inclusion of a resight are common. Based on our findings, when population growth rate is of interest, particularly if preliminary estimates are close to 1, it is advisable to explore the extent of potential bias for the relevant parameter space through simulation, and identify and include precaptures in the catalog where warranted and feasible (e.g., small catalog or automated matching). In data-limited situations, such as that reported on here, additional records may add valuable sample size as well. Precaptures should be excluded from data for abundance

estimation.

The results of mark-recapture analysis of 11 years of photo-ID data for *Z. cavirostris* in a sonar-rich Navy training range indicate long-term site fidelity and high apparent annual survival, but did not provide conclusive evidence regarding trend in abundance. With continued monitoring, the enhanced capture probability from the collaboration between acoustic and field teams—and the site fidelity of *Z. cavirostris* to SOAR—provide an opportunity to link documented short-term behavioral responses of *Z. cavirostris* to MFAS in SOAR to cumulative effects on vital rates and to trends in abundance, and to validate predictions of population-level impacts from Population Consequences of Disturbance models (Fleishman, Costa, Kraus, Moretti, New, & wells, 2016; National Research Council, 2005; New et al., 2013; Pirotta et al., 2018). Achieving this goal depends on continuing this rare, long photo-ID time series while collecting complementary data on vital rates of individuals using SOAR, both high research priorities for research on impacts of MFAS on beaked whales (de Quirós et al., 2019), as well as further development of analytical methods.

DATA ACCESSIBILITY

Sightings and effort data and R code for data processing and analysis are available from the Dryad repository (Curtis et al., 2020).

ACKNOWLEDGMENTS

We thank the U.S. Navy, including N45, the Living Marine Resources program, and the U.S. Navy Pacific Fleet, for funding fieldwork and photo-ID analysis. We thank all our collaborators from the Naval Undersea Warfare Center M3R group including Stephanie Watwood, Susan Jarvis, Ron Morrissey, Nancy DiMarzio, and Karin Dolan, as well as Cascadia Research Collective. The staff of the Southern California Offshore Range and personnel on San Clemente Island provided essential logistical support for field work. NOAA Southwest Fisheries Science Center provided workspace and resources for analysis, and Rich Cosgrove assisted with use of the SWFSC computing cluster. Several NOAA colleagues provided helpful feedback and insights during analysis and writing, including Mike Kinney, Brett McClintock, Jeff Laake, and Jim Carretta. We also thank two internal reviewers, Mike Kinney and Doug Kinzey, and two anonymous reviewers for their

helpful comments on the manuscript. All field work was conducted under NOAA Office of Protected Resources Permits 540-1811, 16111, and 15330.

ANIMAL ETHICS

All field work was covered under an IACUC issued to Cascadia Research Collective. Two of the authors (E.A.F. and G.S.S.) are co-investigators on all named permits for this study. Work under these permits was approved by Cascadia's IACUC and performed in accordance with their Standard Operating Procedures.

FUNDING

K.A.C., E.A.F., and G.S.S. were supported by Office of Naval Research Award Number N0001415IP00088. Collection and processing of photo-ID data by E.A.F., G.S.S., E.L.K., and D.J.M. were supported by awards from the U.S. Navy's Living Marine Resources program, the U.S. Navy's N45 program, and the U.S. Navy's Pacific Fleet. K.A.C. was also supported by a contract from the NOAA Marine Mammal and Turtle Division/Southwest Fisheries Science Center to Ocean Associates, Inc.

COMPETING INTERESTS

E.A.F., G.S.S., and D.J.M. have received additional funding from

the U.S. Navy Living Marine Resources program, Office of Naval Research, and Pacific Fleet for further research related to the work reported here. D.J.M. is a former (retired) employee of the U.S. Navy.

AUTHOR CONTRIBUTIONS

E.A.F., G.S.S., J.E.M., D.J.M., and J.P.B. conceived of the study. E.A.F., G.S.S., and D.J.M. designed the field data collection study and conducted all on-water data collection. E.L.K. and E.A.F. conducted all photo-ID and associated data processing, and combined and proofed all data streams. K.A.C. helped proof the data, designed and implemented the data analysis, drafted the initial text, and revised the manuscript. E.A.F., G.S.S., and E.L.K. advised on data analysis, helped draft the initial text, and assisted in revising the manuscript. J.E.M. assisted in model development and interpretation and assisted in revising the manuscript. D.J.M. led the design and maintenance of the M3R system, coordinated SOAR access, conducted real-time acoustic monitoring during field efforts, and assisted in revising the manuscript. J.P.B. assisted in data interpretation and assisted in revising the manuscript. All

authors contributed to study design and implementation, approved the final submission, and agree to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

REFERENCES

Allen, B. M., Brownell, R. L., Yamada, T. K., & Mead, J. G. (2012). Review of current knowledge on *Ziphius cavirostris* in the North Pacific and North Indian oceans, including identification of knowledge gaps and suggestions for future research. Paper SC/64/SM34 presented to the International Whaling Commission, Scientific Committee, Panama City, Panama.

Barlow, J. (2015). Inferring trackline detection probabilities, *g*(0), for cetaceans from apparent densities in different survey conditions. *Marine Mammal Science*, *31*, 923–943.

Baumann-Pickering, S., Hildebrand, J. A., Yack, T., & Moore, J. E. (2015). *Modeling of habitat and foraging behavior of beaked whales in the Southern California Bight* (Final Report). Office of Naval Research Award Number N00014-12-1-

0273.

Baumann-Pickering, S., Roch, M. A., Brownell, R. L., Jr., Simonis, A. E., McDonald, M. A., Solsona-Berga, A., … Hildebrand, J. A. (2014). Spatio-temporal patterns of beaked whale echolocation signals in the North Pacific. *PLoS ONE*, *9*(3), e86072.

Bonner, S. J., & Holmberg, J. (2013). Mark-recapture with multiple, non-invasive marks. *Biometrics*, *69*, 766–775.

Boulanger, J., McLellan, B. N., Woods, J. G., Proctor, M. F., & Strobeck, C. (2004). Sampling design and bias in DNA-based capture-mark-recapture population and density estimates of grizzly bears. *Journal of Wildlife Management*, *68*, 457–469.

Cañadas, A., Aguilar de Soto, N., Aissi, M., Arcangeli, A., Azzolin, M., B-Nagy, A., … Roger, T. (2018). The challenge of habitat modelling for threatened low density species using heterogeneous data: The case of Cuvier’s beaked whales in the Mediterranean. *Ecological Indicators*, *85*(Supplement C), 128–136.

Carothers, A. D. (1973). The effects of unequal catchability on Jolly-Seber estimates. *Biometrics*, *29*, 79–100.

Claridge, D. E. (2013). *Population ecology of Blainville's beaked whales* (Mesoplodon densirostris) (Ph.D. thesis). University of St. Andrews, St. Andrews, Scotland.

Cooch, E. G., & White, G. C. (2016). Appendix F: Parameter identifiability by data cloning. In E. G. Cooch & G. C. White (Eds.) *Program Mark: A gentle introduction* (pp. F1–F21). Ithaca, NY: Cornell University.

Coomber, F., Moulins, A., Tepsich, P., & Rosso, M. (2016). Sexing free-ranging adult Cuvier's beaked whales (*Ziphius cavirostris*) using natural marking thresholds and pigmentation patterns. *Journal of Mammalogy*, *97*, 879–890.

Cormack, R. M. (1964). Estimates of survival from the sighting of marked animals. *Biometrika*, *51*, 429–438.

Cormack, R. M. (1972). The logic of capture-recapture estimates. *Biometrics*, *28*, 337–343.

Cox, T. M., Ragen, T. J., Read, A. J., Vos, E., Baird, R. W., Balcomb, K., … Benner, L. (2006). Understanding the impacts of anthropogenic sound on beaked whales. *Journal of Cetacean Research and Management*, *7*, 177–187.

Curtis, S. M. (2015). mcmcplots: Create plots from MCMC output.

R package version 0.4.2. https://CRAN.R-project.org/package=mcmcplots

Curtis, K. A., Falcone, E. A., Schorr, G. S., Moore, J. E., Moretti, D. J., Barlow, J., & Keene, E. (2020). Abundance, survival, and annual rate of change in Cuvier's beaked whales (*Ziphius cavirostris*) on a Navy sonar range [Data set]. Dryad. https://doi.org/10.5061/dryad.547d7wm4j

D'Amico, A., Gisner, R. C., Ketten, D. R., Hammock, J. A., Johnson, C., Tyack, P. L., & Mead, J. (2009). Beaked whale strandings and naval exercises. *Aquatic Mammals*, *34*, 452–472.

de Quirós, Y. B., Fernandez, A., Baird, R. W., Brownell, R. L., Jr., de Soto, N. A., Allen, D., … Schorr, G. (2019). Advances in research on the impacts of anti-submarine sonar on beaked whales. *Proceedings of the Royal Society B: Biological Sciences*, *286*, 20182533.

DeRuiter, S. L., Southall, B. L., Calambokidis, J., Zimmer, W. M. X., Sadykova, D., Falcone, E. A., … Tyack, P. L. (2013). First direct measurements of behavioural responses by Cuvier's beaked whales to mid-frequency active sonar.

Biology Letters, *9*, 20130223.

DiMarzio, N., Jones, B., Moretti, D., Thomas, L., & Oedekoven, C. (2018). *Marine mammal monitoring on navy ranges (M3R) on the Southern California Offshore Range (SOAR) and the Pacific Missile Range Facility (PMRF) 2017* (Combined Pacific Annual Monitoring Report [HSTT, MITT, NWTT, GOA TMAA]). Newport, RI: Naval Undersea Warfare Center.

Dorazio, R. M., & J. A. Royle. (2003). Mixture models for estimating the size of a closed population when capture rates vary among individuals. *Biometrics*, *59*, 351–364.

Faerber, M. M., & Baird, R. W. (2010). Does a lack of observed beaked whale strandings in military exercise areas mean no impacts have occurred? A comparison of stranding and detection probabilities in the Canary and main Hawaiian Islands. *Marine Mammal Science*, *26*, 602–613.

Falcone, E. A., & Schorr, G. S. (2014). *Distribution and demographics of marine mammals in SOCAL through photo-identification, genetics, and satellite telemetry* (Report prepared for the Chief of Naval Operations, Washington DC). Published by the Naval Postgraduate School, Monterey, CA

92943 under NPS Grant N00244-10-1-0050. http://calhoun.nps.edu/handle/10945/44226

Falcone, E. A., Schorr, G. S., Douglas, A. B., Calambokidis, J., Henderson, E., McKenna, M. F., … Moretti, D. (2009). Sighting characteristics and photo-identification of Cuvier's beaked whales (*Ziphius cavirostris*) near San Clemente Island, California: A key area for beaked whales and the military? *Marine Biology*, *156*, 2631–2640.

Falcone, E. A., Schorr, G. S., Watwood, S. L., DeRuiter, S. L., Zerbini, A. N., Andrews, R. D., … Moretti, D. J. (2017). Diving behaviour of Cuvier's beaked whales exposed to two types of military sonar. *Royal Society Open Science*, *4*, 170629.

Filadelfo, R., Mintz, J., Michlovich, E., D'Amico, A., Tyack, P. L., & Ketten, D. R. (2009a). Correlating military sonar use with beaked whale mass strandings: What do the historical data show? *Aquatic Mammals*, *35*, 435–444.

Filadelfo, R., Pinelis, Y. K., Davis, S., Chase, R., Mintz, J., Wolfanger, J., … D'Amico, A. (2009b). Correlating whale strandings with navy exercises off Southern California.

Aquatic Mammals, *35*, 445–451.

Fleishman, E., Costa, D. P., Kraus, S., Moretti, D., New, L. F., & Wells, R. S. (2016). Monitoring population level responses of marine mammals to human activities. *Marine Mammal Science*, *32*, 1004–1021.

Fleming, A.H., Kellar, N. M., Allen, C. D., & Kurle, C. M. (2018). The utility of combining stable isotope and hormone analyses for marine megafauna research. *Frontiers in Marine Science*, *5*, 338.

Gimenez, O., Cam, E., & Gaillard, J.-M. (2018). Individual heterogeneity and capture-recapture models: What, why and how? *Oikos*, *127*, 664–686.

Gupta, M., Joshi, A., & Vidya, T. N. (2017). Effects of social organization, trap arrangement and density, sampling scale, and population density on bias in population size estimation using some common mark-recapture estimators. *PloS ONE*, *12*(3), e0173609.

Harwood J., King, S. Booth, C. Donovan, C. Schick, R. S. Thomas, L., & New, L. (2016) Understanding the population consequences of acoustic disturbance for marine mammals. In

A. Popper & A. Hawkins (Eds.), *The effects of noise on aquatic life II. Advances in experimental medicine and biology* (Vol. 875). New York, NY: Springer.

Hilborn, R., & Walters, C. J. (1992). *Quantitative fisheries stock assessment, choice, dynamics and uncertainty*. London, UK: Chapman and Hall.

Hines, J. E., & Nichols, J. D. (2002). Investigations of potential bias in the estimation of λ using Pradel's (1996) model for capture-recapture data. *Journal of Applied Statistics*, *29*, 573-587.

Huggins, R. M. (1989). On the statistical analysis of capture experiments. *Biometrika*, *76*, 133–140.

Jarvis, S. M., Morrissey, R. P., Moretti, D. J., DiMarzio, N. A., & Shaffer, J. A. (2014) Marine mammal monitoring on navy ranges (M3R): A toolset for automated detection, localization, and monitoring of marine mammals in open ocean environments. *Marine Technology Society Journal*, *48*, 5–20.

Jolly, G. M. (1965). Explicit estimates from capture–recapture data with both death and immigration-stochastic model.

Biometrika, *52*, 225–247.

Lele, S. R., Dennis, B., & Lutscher, F. (2007). Data cloning: Easy maximum likelihood estimation for complex ecological models using Bayesian Markov chain Monte Carlo methods. *Ecology Letters*, *10*, 551–563.

Lele, S. R., Nadeem, K., & Schmuland, B. (2010). Estimability and likelihood inference for generalized linear mixed models using data cloning. *Journal of the American Statistical Association*, *105*, 1617–1625.

Link, W.A., Yoshizaki, J., Bailey, L. L., & Pollock, K. H. (2010). Uncovering a latent multinomial: Analysis of mark-recapture data with misidentification. *Biometrics*, *66*, 178–185.

MacKenzie, D. I., Nichols, J. D., Sutton, N., Kawanishi, K. & Bailey, L. L. (2005), Improving inferences in population studies of rare species that are detected imperfectly. *Ecology*, *86*, 1101–1113.

McCarthy, E., Moretti, D., Thomas, L., DiMarzio, N., Morrissey, R., Jarvis, … Dilley, A. (2011). Changes in spatial and temporal distribution and vocal behavior of Blainville's

beaked whales (*Mesoplodon densirostris*) during multiship exercises with mid-frequency sonar. *Marine Mammal Science*, *27*, E206–E226.

McClintock, B. (2015). *multimark*: An R package for analysis of capture recapture data consisting of multiple “noninvasive” marks. *Ecology and Evolution*, *5*, 4920–4931.

McClintock, B. T., Conn, P. B., Alonso, R. S., & Crooks, K. R. (2013). Integrated modeling of bilateral photo-identification data in mark-recapture analyses. *Ecology*, *94*, 1464–1471.

McSweeney, D. J., Baird, R. W., & Mahaffy, S. D. (2007). Site fidelity, associations, and movements of Cuvier’s (*Ziphius cavirostris*) and Blainville’s (*Mesoplodon densirostris*) beaked whales off the island of Hawai‘i. *Marine Mammal Science*, *23*, 666–687.

Moore, J. E., & Barlow, J. P. (2013). Declining abundance of beaked whales (Family Ziphiidae) in the California Current Large Marine Ecosystem. *PLoS ONE*, *8*(1), e52770.

Moore, J. E. & Barlow, J. P. (2017). *Population abundance and trend estimates for beaked whales and sperm whales in the*

California Current from ship-based visual line-transect survey data, 1991–2014 (NOAA Technical Memorandum NOAA-TM-NMFS-SWFSC-585). Washington, DC: U.S. Department of Commerce.

Moretti, D. (2014). *Marine mammal monitoring on Navy ranges 2014* (Summary report). https://www.navymarinespeciesmonitoring.us/files/1614/2767/1131/Moretti_2015_M3R_2014_Summary_Report_SCORE_Feb2015.pdf

Morrissey, R. P., Ward, J., DiMarzio, N., Jarvis, S., & Moretti, D. J. (2006). Passive acoustic detection and localization of sperm whales (*Physeter macrocephalus*) in the Tongue of the Ocean. *Applied Acoustics*, *67*, 1091–1105.

National Academy of Sciences, Engineering, and Medicine. (2017). *Approaches to understanding the cumulative effects of stressors on marine mammals*. Washington, DC: The National Academies Press.

National Research Council. (2003). *Ocean noise and marine mammals*. Washington, DC: National Academy Press.

National Research Council. (2005). *Marine mammal populations and ocean noise: Determining when noise causes biologically*

significant effects. Washington, DC: National Academy Press.

New, L. F., Moretti, D. J., Hooker, S. K., Costa, D. P., & Simmons, S. E. (2013). Using energetic models to investigate the survival and reproduction of beaked whales (Family Ziphiidae). *PLoS ONE*, *8*(7), e68725.

Nowacek, D. P., Thorne, L. H., Johnston, D. W. & Tyack, P. L. (2007). Reponses of cetaceans to anthropogenic noise. *Mammal Review*, *37*, 81–115.

Pace, R. M., III, Corkeron, P. J., & Kraus, S. D. (2017). State–space mark–recapture estimates reveal a recent decline in abundance of North Atlantic right whales. *Ecology and Evolution*, *7*, 8730–8741.

Pirotta, E., Booth, C. G., Costa, D. P., Fleishman, E., Kraus, S. D., Lusseau, D., … Harwood, J. (2018). Understanding the population consequences of disturbance. *Ecology and Evolution*, *8*, 9934–9946.

Plummer, M. (2016). *rjags*: Bayesian Graphical Models using MCMC. R package version 4-6. http://CRAN.R-project.org /package=rjags.

Plummer, M., Best, N., Cowles, K., & Vines, K. (2006). CODA: Convergence diagnosis and output analysis for MCMC, *R News*, *6*, 7–11.

Pradel, R. (1996). Utilization of capture-mark-recapture for the study of recruitment and population growth rate. *Biometrics*, *52*, 703–709.

Pradel, R., Hines, J. E., Lebreton, J. D. & Nichols, J. D. (1997). Capture-recapture survival models taking account of transients. *Biometrics*, *53*, 60–72.

R Core Team. (2016). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Royle, J. A., & Dorazio, R. M. 2008. *Hierarchical modeling and inference in ecology*. London, UK: Academic Press.

Richardson, W. J., Greene, Jr., C. R., Malme, C. I., & Thomson, D. H. (Eds). (1995). *Marine mammals and noise*. San Diego, CA: Academic Press.

Rosso, M., Ballardini, M., Moulins, A., & Würtz, M. (2011). Natural markings of Cuvier's beaked whale *Ziphius cavirostris* in the Mediterranean Sea. *African Journal of*

Marine Science, *33*, 45–57.

Schorr, G. S., Falcone, E. A., Moretti, D. J., & Andrews, R. D. (2014). First long-term behavioral records from Cuvier's beaked whales (*Ziphius cavirostris*) reveal record-breaking dives. *PLoS ONE*, *9*(3), e92633.

Seber, G. A. F. (1965). A note on the multiple recapture census. *Biometrika*, *52*, 249-259.

Steiger, G. H., & Calambokidis, J. (2000). Reproductive rates of humpback whales off California. *Marine Mammal Science*, *16*, 220–239.

Stevick, P. T., Palsbøll, P. J., Smith, T. D., Bravington, M. V. & Hammond, P. S. (2001). Errors in identification using natural markings: Rates, sources, and effects on capture–recapture estimates of abundances. *Canadian Journal of Fisheries and Aquatic Sciences*, *58*, 1861–1870.

Suárez, C. R. (2018). *Abundance estimate, survival and site fidelity patterns of Blainville's* (Mesoplodon densirostris) *and Cuvier's* (Ziphius cavirostris) *beaked whales off El Hierro (Canary Islands)* (M.Phil. thesis). University of St. Andrews, St. Andrews, Scotland.

Taylor, B. L. & T. Gerrodette, T. (1993). The uses of statistical power in conservation biology: The vaquita and Northern Spotted Owl. *Conservation Biology*, *7*, 489–500.

Tenan, S., Pradel, R., Tavecchia, G., Igual, J. M., Sanz-Aguilar, A., Genovart M., & Oro, D. (2014). Hierarchical modelling of population growth rate from individual capture-recapture data. *Methods in Ecology and Evolution*, *5*, 606–614.

Tyack, P. L., Zimmer, W. M. X., Moretti, D., Southall, B. L., Claridge, D. E., Durban, J. W., … Boyd, I. L. (2011) Beaked whales respond to simulated and actual Navy sonar. *PLoS ONE*, *6*(3), e17009.

United States Geological Survey. (2009). Topography, NOAA Coastal Relief Model, 3 arc second, Vol. 6 (Southern California). Download provided by the NOAA ERDDAP from https://coastwatch.pfeg.noaa.gov/erddap/griddap/usgsCeCrm6.html

Whitehead, H., & Gero, S. (2015). Conflicting rates of increase in the sperm whale population of the eastern Caribbean: Positive observed rates do not reflect a healthy

population. *Endangered Species Research*, *27*, 207–218.

Williams, R., Wright, A. J., Ashe, E., Blight, L. K., Bruintjes, R., Canessa, R., … Wale, M. A. (2015). Impacts of anthropogenic noise on marine life: Publication patterns, new discoveries, and future directions in research and management. *Ocean and Coastal Management*, *115*, 17–24.

Wilson, B., Hammond, P. S., & Thompson, P. M. (1999). Estimating size and assessing trends in a coastal bottlenose dolphin population. *Ecological Applications*, *9*, 288–300.

TABLE 1 Capture frequency summaries for annual captures of *Ziphius cavirostris*. Total number of individuals (*n*), average annual capture frequency (*x*), and capture frequency counts for individuals whose capture histories contain photographs (1) from only the left side, (2) from both sides within a sighting or characterized by a distinctively damaged or shaped fin, and (3) from only the right side. Analyses with *multimark* models included all three types of capture histories. The Pradel-lambda analysis was limited to (4) right-sided capture histories from individuals known only from the right side or from both sides.

			Capture frequency						
Capture history type	***n***	***x***	**1**	**2**	**3**	**4**	**5**	**6**	**7**
Left side only	18	1.06	17	1	—	—	—	—	—
Both sides or fin	87	1.69	54	16	11	4	1	—	1
Right side only	21	1.14	18	3	—	—	—	—	—
All right-sided captures	105	1.42	74	22	6	2	1	—	—

TABLE 2 Relative probabilities of the four candidate *multimark* CJS models included in RJMCMC sampling. Model variations for capture probability include intercept-only, intercept plus individual random effects ε_i, and intercept plus ε_i and either β_1, annual hours of search effort in Beaufort sea states 0 through 2, or β_2, effort adjusted for season and Beaufort sea state.

Model	Posterior probability
$p(.)\ \phi(.)$	0.013
$p(\sim 1 + \varepsilon_i)\ \phi(.)$	0.265
$p(\sim 1 + \beta_1 + \varepsilon_i)\ \phi(.)$	0.275
$p(\sim 1 + \beta_2 + \varepsilon_i)\ \phi(.)$	0.447

TABLE 3 Summaries of posterior distributions for ϕ (apparent annual survival), N_{nc} and N (corrected annual abundance excluding calves and corrected total annual abundance, respectively), and λ (population growth rate). P_x is the xth percentile of the posterior distribution for each parameter, *SD* is standard deviation, and n is effective sample size. Estimate of ϕ is from the most probable *multimark* CJS model fitted to capture histories from both sides simultaneously, based on RJMCMC (above). Abundance was estimated from a closed-population *multimark* model fitted to capture histories from both sides from the most recent three occasions, spanning 2015–2018, and multiplied by either a correction factor of 1.052 to account for noncalf *Ziphius* that did not meet distinctiveness criteria (N_{nc}) or 1.107 to account for both insufficiently marked animals and calves (N). λ was estimated with a Pradel -lambda model fitted to right-side capture histories and imputed over 250 values of ϕ drawn from the posterior distribution from the *multimark* CJS model (see Methods).

Model		Mean	*SD*	Mode	P_5	P_{25}	P_{50}	P_{75}	P_{95}	*n*
multimark CJS	ϕ	0.950	0.027	0.962	0.899	0.934	0.954	0.970	0.986	9,811
multimark Closed	N_{nc}	115	49	84	67	83	101	131	208	5,207
multimark Closed	N	121	52	89	71	87	106	138	219	5,207
Pradel-lambda	λ	0.992	0.029	0.989	0.944	0.972	0.991	1.011	1.041	39,021

FIGURE 1 Locations of *Ziphius cavirostris* photo identifications included in analyses (white triangles), overlaid on bathymetry (United States Geological Survey, 2009). Each location represents a sighting, and so may include several photo identifications. San Clemente Island (source: Global Self-consistent, Hierarchical, High-resolution Shoreline; NOAA National Geophysical Data Center) and 800-m isobaths are delineated in black, and the Southern California Anti-submarine Warfare Range (SOAR) is outlined in white. Inset shows location of study area within the Southern California Bight.

FIGURE 2 Workflow for mark-recapture analysis. Filtering level for each step is indicated on the left (gray bubbles), filter criterion on the right. Precaptures are photos of catalog-qualifying individuals taken before the catalog-qualifying photo was taken. Technically, precaptures for individuals in the mark-recapture data set were added retroactively.

*Although simulations showed that including precaptures could lead, on average, to slight positive bias in survival estimates (Appendix S1), preliminary analysis of the data set at hand resulted in equal or lower survival estimates when included, so

they were included.

FIGURE 3 Total hours of search effort for *Ziphius cavirostris* in the San Nicolas Basin by season per annual occasion. Gray represents “summer” effort, defined as effort conducted during July through October, based on observed shifts in vocalization detection frequency of *Z. cavirostris* from passive acoustic studies in the Southern California Bight (Baumann-Pickering et al., 2015). Black represents nonsummer effort (all other months). Annual occasions begin in August of the occasion year. For periods when two boats were simultaneously on effort (during 2007–2008), individual vessel efforts were summed.

FIGURE 4 Capture rates with Beaufort Sea State and season for *Ziphius cavirostris* in the San Nicolas Basin (based on 11 years of photo-ID data). Captures are defined as sighting-level recapture-qualifying records, regardless of inclusion in the data set for analysis (i.e., not filtered for individuals with catalog-qualifying records, by age class, or by distinctiveness), from either left- or right-side encounters. Gray triangles represent summer (S) capture rates, defined as July through October; black circles represent nonsummer (NS)

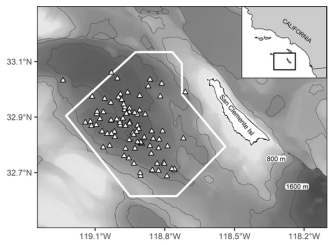
rates.

FIGURE 5 Discovery curves for right-side (black, circles) and left-side (gray, triangles) capture histories of total individuals versus number of identifications made on an annual basis from August 2007 through July 2018. Dashed line is 1:1.

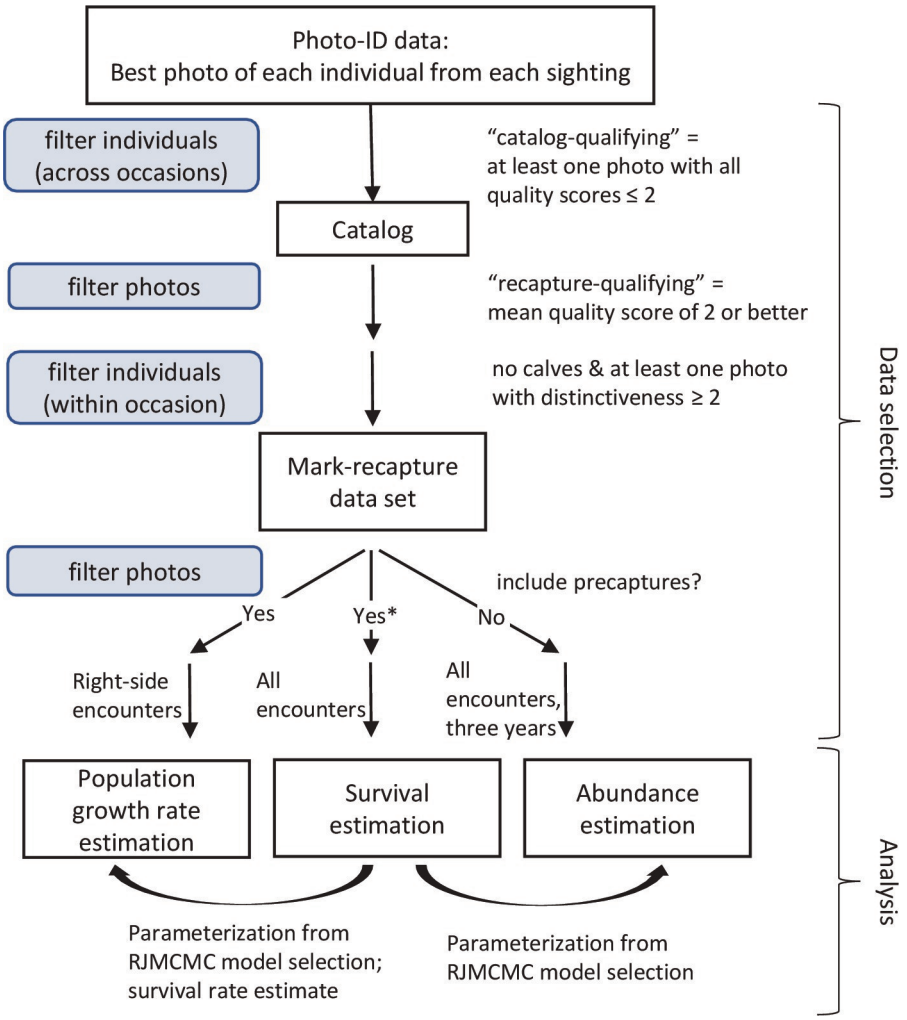
FIGURE 6 Posterior probability densities for estimated parameters for *Ziphius* using SOAR on an ongoing basis. Estimates of apparent annual survival rate ϕ and probit-link slope of annual capture probability with the Beaufort- and season-adjusted effort index are from the most probable *multimark* CJS model fit, based on RJMCMC (above). Total abundance *N* as estimated from a closed-population *multimark* model fitted to data from the most recent three occasions, spanning 2015–2018, and includes a correction factor of 1.107 to account for *Ziphius* that did not meet distinctiveness criteria and calves. Population growth rate λ for *Ziphius* using SOAR was estimated with a Pradel-lambda model fitted to right-side capture histories and imputed over 250 values of ϕ drawn from the posterior distribution from the *multimark* CJS model (see Methods).

FIGURE 7 Overprotective (black) and underprotective (gray) error rates for detecting a decline in abundance at varying levels of quasi-α and rates of decline, given current sampling effort and inference from a Bayesian Pradel-lambda model with fixed and random temporal effects (11 years, capture probability of 0.085, and fixed temporal effects and temporal and individual random effects as specified in Methods). Overprotective error is the probability of falsely inferring a trend if none exists. Underprotective error is the probability of failing to detect a true decline. Error rates for each scenario (line) are based on 1,000 simulations. Quasi-α is the threshold at which evidence is considered "strong" that abundance is declining.

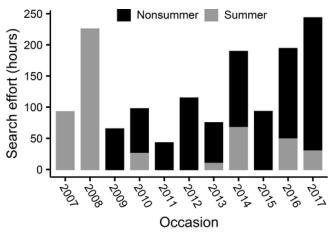
FIGURE 8 Probability of detecting a declining trend with additional years of effort or increased effort per year. Probability of detecting decline is the percentage of simulations with 80% posterior probability of λ (population growth rate) <1, given model convergence. Black points and lines correspond to scenarios with $\lambda = 0.966$; gray points and lines to $\lambda = 0.933$.



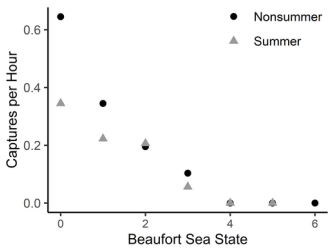
MMS_12747_4940_Fig.1.tiff



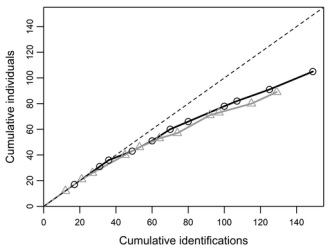
mms_12747_4940_fig.2.eps



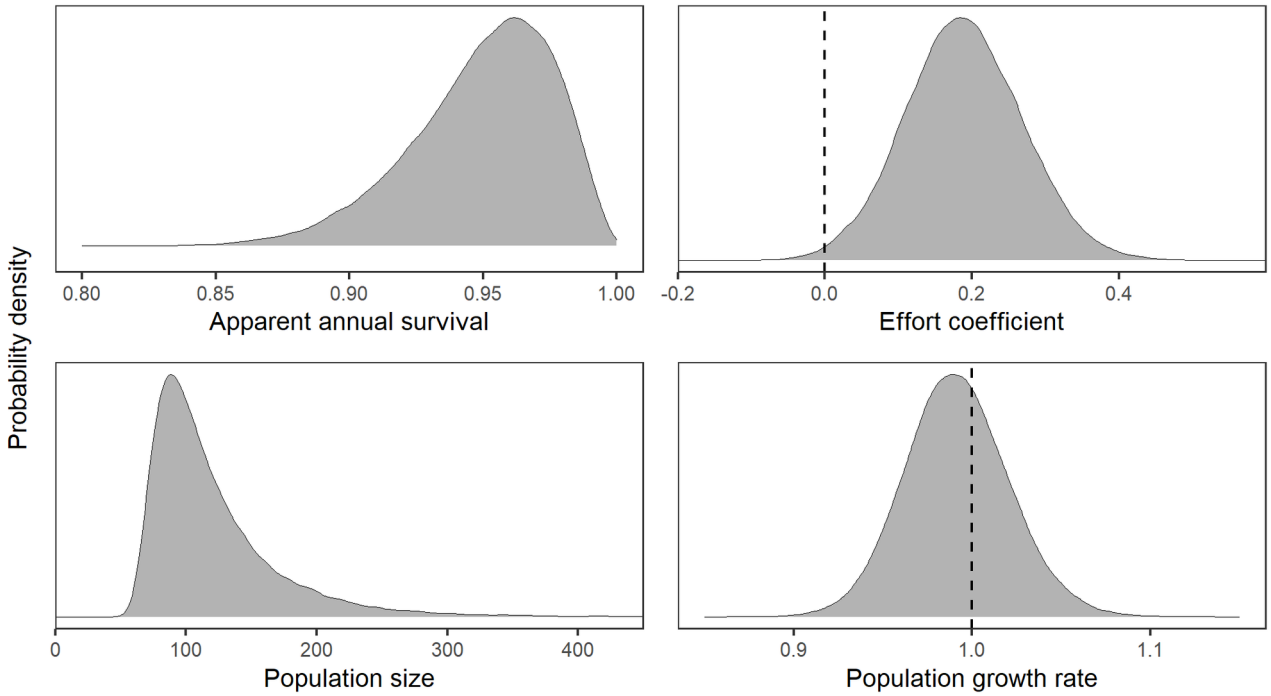
MMS_12747_4940_Fig.3.tiff



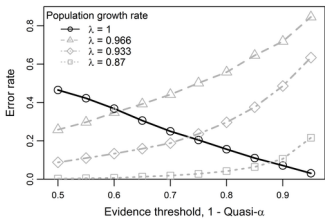
MMS_12747_4940_Fig.4.tiff



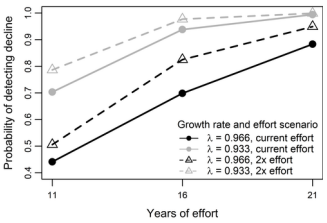
MMS_12747_4940_Fig.5.tiff



MMS_12747_4940_Fig.6.tiff



MMS_12747_4940_Fig.7.tiff



MMS_12747_4940_Fig.8.tiff