Accounting for spatiotemporal sampling bias in a long-term dataset establishes a decline in abundance of endangered false killer whales (*Pseudorca crassidens*) in the main Hawaiian Islands

Keywords: telemetry, data integration, mark-recapture, endangered species, Jolly-Seber, false killer whales

Janelle J. Badger[1,*], Robin W. Baird[2], Devin S. Johnson[1], Amanda L. Bradford[1], Sabre D. Mahaffy[2], Michaela A. Kratofil[2,3,4], Tori Cullins[5], Jens J. Currie[6], Stephanie H. Stack[6], Erin M. Oleson[1]

[1] Pacific Islands Fisheries Science Center, NMFS, NOAA, 1845 Wasp Boulevard, Building 176, Honolulu, HI 96818, USA

[2] Cascadia Research Collective, 218½ West 4th Avenue, Olympia, WA 98501, USA

[3] Marine Mammal Institute, Oregon State University, 2030 SE Marine Science Dr, Newport, Oregon 97365, USA

[4] Department of Fisheries, Wildlife, and Conservation Sciences , Oregon State University, 2820 Campus Way, Corvallis, Oregon 97331, USA

[5] Wild Dolphin Foundation, 87-1002 C Hakimo Place, Waianae, HI, 96792, USA

[6] Pacific Whale Foundation, 300 Ma‘alaea Road, Suite 211, Wailuku, HI 96793, USA

Abstract

We estimated abundance from 1999-2021 for the endangered main Hawaiian Island (MHI) insular population of false killer whales using a modeling technique that incorporates animal availability into a capture-recapture model. Since 1999, this population has been sampled by a variety of dedicated systematic, nonsystematic, and opportunistic methods that have resulted in the yearly encounter histories of 202 individuals and 52 satellite telemetry tag deployments used in this analysis. Survey efforts and animal location data from telemetry deployments were separately analyzed using kernel density estimators, and the degree of overlap between these two processes was used this overlap as a covariate for detection probability within a Jolly-Seber open population model in a Bayesian framework. Using simulated data, we found this model estimates abundance with a higher degree of accuracy and precision than conventional methods and is robust to many sampling and ecological complications, such as variable tag deployment lengths, low detectability, and variable social group sizes. Upon fitting the model to the false killer whale dataset, we found that the insular population of false killer whales remains small, with an estimated 138 individuals (95% CRI: [120, 160]) in the population in 2021. The population appears to be in decline throughout the study period, with an estimated trend of -3.51 (95% CRI: [-8.40, 2.04]) over the entire time series and -5.53, 95% CRI = [-9.91, -1.61] over the past 10 years. As this modeling approach addresses the spatiotemporal variability in sampling effort, these updated abundance estimates provide improved inputs for management frameworks.

Introduction

The importance of detecting declines in wildlife populations has long been recognized. Even in cases where species extinction is not imminent, population declines can lead to extirpations and losses in ecosystem function and services (Ceballos and Ehrlich 2002). Identifying that population declines are occurring and the factors that may be driving losses of individuals are needed to inform conservation and management planning. Detecting trends in population growth, positive or negative, allows for an evaluation of the efficacy of management actions and can inform future decision-making (Yuccoz et al. 2001). However, the challenges of data collection, including for dispersed and inaccessible species such as marine mammals, can lead to time series of abundance or related indices that lack the statistical power to detect declines (e.g., Taylor et al. 2007). This predicament emphasizes the importance of incorporating novel analytical techniques and auxiliary datasets with what may otherwise be considered the best available science (e.g., Murphy and Weiland 2016) when assessing marine mammal and other difficult-to-study wildlife populations.

False killer whales (Pseudorca crassidens) are an example of such species whose life history and behavior makes them challenging to survey, and consequently collect data necessary for robust estimates of population dynamics. False killer whales are long-lived (well into their 60's) and slow to mature (10-15 years old at sexual maturity; Ferreira et al. 2014; Photopoulou et al. 2017). They are strongly social, known to travel in large groups and exhibit cohesive social structure (Baird et al. 2008; Mahaffy et al. 2023). They inhabit sub-tropical and tropical oceanic domains worldwide (Baird, 2008; Zaeschmar and Estrela, 2022), and a number of coastal and island-associated populations have been documented (Baird et al. 2008, Baird, 2016, Silva et al, 2014, Palmer et al. 2017, Douglas et al. 2023, Zaeschmar 2014). Their generally offshore distribution, and their tendency to move frequently over large spatial domains (Baird et al. 2012), makes it challenging to adequately survey populations.

The main Hawaiian islands (MHI) insular population of false killer whales is small, last estimated to number 167 ± 23 individuals in 2015 (Bradford et al. 2018). Individuals in this population may move widely among and frequently between island areas within the main Hawaiian Islands (Baird et al. 2012), making them sparse throughout their range. This population is known to preferentially associate in social groups, hereafter referred to as "clusters" (Mahaffy et al. 2023). Four stable social groups have been recognized (Mahaffy et al. 2023) and consist of family members and

regular associates (Martien et al. 2019). There has been some evidence for cluster-specific space use patterns (Baird et al. 2012, Baird et al. 2023). When encountered during survey effort, individuals and subgroups within clusters are often spread out, traveling kilometers apart (Bradford et al. 2014, Baird 2016). The MHI population of false killer whales has declined in recent decades and was listed as endangered under the Endangered Species Act in 2012 (Oleson et al. 2010). The greatest suspected threats to this population's viability include interactions with nearshore fisheries (Baird et al. 2015, 2017), exposure to pollutants (Bachman et al. 2014; Kratofil et al. 2020; Ylitalo et al. 2009), and reduced genetic diversity (Chivers et al. 2010; Martien et al. 2014).

Numerous boat-based surveys, photo-identification, satellite-telemetry, and genetic studies have made the MHI insular population the world's most thoroughly studied population of false killer whales (Baird 2016). However, each of these data streams present unique difficulties to estimate necessary metrics to monitor this endangered population, such as abundance and population growth rate. For example, there is variable boat-based survey effort around the main Hawaiian Islands, and as this population is wide-ranging, these surveys will only encompass a small proportion of the population's range in any given period. Weather and sea conditions generally further restrict areas viable for visual sampling on a smaller scale, with surveys almost exclusively conducted on leeward sides of islands protected from trade winds (Baird et al. 2013, 2024). Without information on individual space use (e.g. if animal movement is random with respect to island geography), it is unclear how this may affect availability to capture and resulting metrics.

Substantial work has been done to estimate abundance for this population, generating estimates robust to many forms of sampling variability and bias (Bradford et al. 2018). However, availability bias is inadequately accounted for within conventional capture-recapture (CR) models. In using conventional CR models, resulting abundance estimates truly only refer to the sampled population in each year, rather than the full population abundance. As it is unclear what proportion of the population is sampled in each year, these estimates are insufficient to determine population trend and thus difficult to incorporate into recovery plans.

Here, we estimate yearly abundance and long-term trend of the m

Methods

Data Collection

Data used in this analysis are predominantly sourced from dedicated nonrandom, nonsystematic small-boat surveys for odontocetes conducted by Cascadia Research Collective (CRC) from 1999-2021 (Figure 1). While CRC has undertaken surveys off all island areas within the MHI, not every area is surveyed every year, with more than half of the effort undertaken off of Hawai'i Island (Baird et al. 2024). Generally, between 1-6 occasions of such efforts lasting 1-6 weeks were conducted throughout each year, with surveyed areas designed to maximize likelihood of animal encounter (details of the field operations are provided in Baird et al. (2024). Even with this focused sampling, false killer whales are only encountered in just over 3% of such surveys. At each false killer whale group sighting, researchers recorded the location of the group and took photographs for individual identification based on the prevalence of permanent markings (e.g. nicks, notches) on the leading and trailing edge of the dorsal fin(Baird et al. 2008). Individuals considered "distinctive" or "very distinctive" (Baird et al. 2008), were assigned to 1 of 4 identified social clusters using the analysis of network modularity explained in Mahaffy et al (in review).

Along with their own archive, CRC has curated photos from other research groups, such as NOAA's Pacific Islands Fisheries Science Center (PIFSC), the Pacific Whale Foundation (PWF), and Wild Dolphin Foundation (WDF), and from ocean users such as whale watch operators and photographers (Figure 1). Encounter data were then compiled at the annual scale, with individuals recorded as encountered or not encountered each year.

When crew expertise, ocean conditions, and animal behavior allow, CRC and PIFSC research efforts also included satellite tag deployments. Between 2007 and 2022, whales were tagged using location-only satellite tags (SPOT5 or SPOT6, Wildlife Computers, Redmond, WA), or location-and-dive behavior-transmitting satellite tags (SPLASH10 or SPLASH10-F (Fastloc®-GPS) Wildlife Computers) in the Low-Impact Minimally-Percutaneous External-electronics Transmitter (LIMPET) configuration (Andrews et al. 2008). Telemetry tagging was undertaken under relevant permits from the National Marine Fisheries Service (NMFS) and tagging methods were approved by the Institutional Animal Care and Use Committees of CRC and PIFSC. Tags were deployed with a pneumatic projector and were attached to the dorsal fin or just below the fin of the whales by two 6.7 cm titanium darts with backward facing petals. Tags were duty-cycled to only transmit during the hours with the highest probability of a satellite being overhead. Tag transmission schedules varied by tag type and by year.

Prior to analyses, location data were filtered through the Douglas-Argos Filter (Douglas et al. 2012; via Movebank, Kranstauber et al. 2011) to remove unrealistic locations based on traveling speeds and turning angles, using settings defined in Baird et al. (2012). Fastloc-GPS locations for relevant tag deployments (n = 2) were filtered by removing locations with residual values greater than 35 and time errors greater than 10 seconds (Dujon et al. 2014) and subsequently running the data through a general speed filter via Movebank (Kranstauber et al. 2011). When applicable, one of each pair of tagged individuals moving in concert were removed prior to analyses to reduce pseudoreplication (see Schorr et al. (2010) for details).

Statistical Analysis

To account for the spatiotemporal variability in sampled areas, we used a novel pseudospatial technique following Badger et al. (in prep). This method involves estimating the coverage of survey effort relative to animal space use, hereafter referred to as *overlap*, and using this variable to inform animal availability within a capture-recapture model. The multistep process is as follows: 1) computing a kernel density estimate of individual tagged whales and scaling to social clusters to estimate population-level *animal space use*; 2) computing kernel density estimates for survey effort from research groups and ocean users to obtain estimates of survey *coverage*; 3) determining availability of animals to survey coverage by finding the interaction between (1) and (2), or the *overlap*, using Bhattacharyya's affinity; and 4) incorporating this overlap measure within the detection process of an open-population Jolly-Seber mark recapture model (Jolly 1965, Seber 1965) fit in a Bayesian framework.

Describing animal space use

First, for each telemetered false killer whale, locations are fit to a continuous-time correlated random walk (CTCRW) model using R package `crawl` (v. 2.2.1, Johnson et al. 2008) to account for location error and predict locations (paths) from observed animal locations. These imputed paths were then rerouted around the islands using the R package `pathroutR` (London 2020). Then, individual utilization distributions (UDs) are described using a kernel density estimator using the R package `kd`. For the KDE estimator, we used a simple plug-in band-width equal to,

$$h_i = n_{ei}^{-1/3} \cdot \Sigma_i,$$

where $\mathbf{\Sigma}_i$ is the covariance matrix among locations for the ith individual and n_{ei} is the effective sample size (ESS), as described by Gelman et al. (2013) and updated in Vehtari et al. (2021), of the ith telemetry data set. The ESS uses the correlation structure of the CTCRW model to assess the effective number of observations that will be less than the total number of observed locations, which will inflate the kernel size. This approach is similar to the autocorrelated kernel density estimate of Fleming et al. (2015) in that it seeks to produce a more predictive UD that accounts for the limited time observation of an animal's correlated travel. The full form of the kernel density estimator for the ith telemtered individual, $\hat{f}_i$, is given below.

Let $(s_1, ..., s_n)$ be telemetry locations, such that $s_j = \{x_j, y_j\}$ where $j \in \{1, ..., n_i\}$ and n_i is the total number of locations for the ith individual. Then,

$$\hat{f}_i = \frac{1}{n_i \cdot h_i} \sum_{j=1}^{n_i} K\left(\frac{s - s_j}{h_i}\right),$$

where K is a Gaussian kernel, i.e. $K(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2}$. The resulting f_i were then normalized to sum to 1 over the study area.

The resulting UD provides a density surface for individual presence in space, however, in order to use the telemetry data information for non-telemetered animals in the CR population, we require an estimate of *population*-level space use to determine animal availability to our survey efforts. There are many ways we might weight individual UDs when averaging (see Conn et al. (2022)); however, we aim to develop an overall use index for use as a covariate in a CR model that can be adjusted to best model detections using the associated coefficients of the model. Therefore, we use the straight average UD. Because this population is known to be affiliated as stable, distinct social clusters, these individual UDs are summed by social cluster to a "cluster UD", henceforth referred to as UD_c, where $c \in \{1, 2, 3, 4\}$ for the current 4 cluster designations (Figure 2).

Defining survey effort

Survey effort for the purposes of this analysis was computed using time-specific kernel density estimates based on contributor coverage of the waters around the MHI. CRC's surveys were nonsystematic and nonrandom, though did attempt to cover a broad range of habitats and space over the course of each sampling period, whereas PWF generally used a systematic approach, and other ocean users opportunistically encountered false killer whales on routine whale- or dolphin-watching

tours or fishing routes. PIFSC encountered false killer whales while on systematic surveys, such as the Hawaiian Islands Cetacean and Ecosystem Assessment Survey (HICEAS, Yano et al. 2018), as well as on nonsystematic and nonrandom survey designs, such as in small-boat efforts searching for cetaceans off the west side of the island of O‘ahu.

As each contributor to the photo-identification archive used varying methods to encounter false killer whales throughout the study, we utilized a flexible approach to defining effort. For three of the four largest contributors to the archive, CRC, PWF, and PIFSC, survey tracks were recorded via an affixed GPS for almost the entire time series. WDF, another main contributor, provided GPS tracks when available, but for many years these were not recorded. WDF along with other organizations and ocean users that frequently contributed photos largely were able to define regularly covered areas that, for many, were visited daily throughout the year. We sampled 1000 points within these defined effort areas for inclusion in the kernel density estimate for each year the contributor was active. Photographs from groups not able to provide any effort information were not included in this analysis, although it should be noted these accounted for only a very small proportion of the total yearly records. We compute survey effort coverage using a simple kernel density estimate of survey tracks and sampled points from effort areas from each year, subsequently referred to as $UD_{e,t}$ for $t \in \{1, ..., T\}$, where T is the data time series length, and with a bandwidth based on a reasonable maximum sighting distance from survey vessels (2 km).

Computing overlap metric

To calculate the overlap between animal space use UD_c and survey efforts for each year $UD_{e,t}$, we computed Bhattacharyya's affinity BA:

$$BA_{c,e,t} = \int_x \int_y \sqrt{UD_c(x,y)} \cdot \sqrt{UD_{e,t}(x,y)},$$

where $UD_c(x,y), UD_{e,t}(x,y)$ are the values of the UD of the animals and survey efforts, respectively, at the point (x,y) in the year t, where $t \in \{1, ..., T\}$. This overlap measure is then standardized over time, i.e. $overlap_{c,t} = \frac{BA_{c,e,t} - \mu}{\sigma}$ where $\mu = \frac{1}{T}\sum_1^T BA_{c,e,t}$ and $\sigma = \sqrt{\frac{\Sigma(BA_{c,e,t} - \mu)^2}{T}}$.

Model form

We use the overlap variable as a covariate in the detection process of a hierarchical formulation of the Jolly-Seber (JS) open population mark-recapture model to estimate abundance (Jolly 1965,

Seber 1965, see Kéry and Schaub 2012). At each sampling occasion (in this case, year), individuals can be in one of three possible states: "not yet entered", "alive", and "dead", and transitions among these states is governed by two ecological processes: entry and survival.

Suppose we have an augmented population of M individuals, of which N are genuine and $M-N$ are pseudo-individuals. Entry is described using ψ, the removal entry probability, i.e. the probability that governs movement from state "not yet entered" to the state "alive". The state of individual i at the first occasion is determined by a Bernoulli trial with probability ψ at the first time step:

$$z_{i,1} \sim \text{Bernoulli}(\psi_1)$$

And subsequent states are determined either by survival ϕ, for an individual already entered ($z_{i,t} = 2$) or by entry, ψ_t, for those that have not ($z_{i,t} = 1$) via a transition matrix:

$$z_{i,t}|z_{i,t-1} \sim \begin{pmatrix} & \text{not yet entered} & \text{alive} & \text{dead} \\ \text{not yet entered} & 1-\psi_t & 0 & 0 \\ \text{alive} & \psi_t & \phi & 0 \\ \text{dead} & 0 & 1-\phi & 1 \end{pmatrix}$$

Similarly, observations are governed by a detection process, p, that depends on an individual's cluster assignment c_i:

$$y_{i,t}|z_{i,t} \sim \begin{pmatrix} & \text{not yet entered} & \text{alive} & \text{dead} \\ \text{observed} & 0 & p_{i,t} & 0 \\ \text{unobserved} & 1 & 1-p_{i,t} & 1 \end{pmatrix}$$

$$\text{Where } \text{logit}(p_{i,t}) = \mu_t + \delta \cdot \text{overlap}_{c_i,t}$$

$$\text{with intercepts } \mu_t, \text{ and } \mu_1 = \mu_2 \text{ for identifiability.}$$

Then, population abundance in a given year N_t is defined as the number of individuals in the "alive" state, that is $N_t = \sum_i \mathbb{I}(z_{i,t} - 2)$, where $\mathbb{I}(x)$ is an indicator that $x = 0$.

To determine trend in abundance for this population, we regressed the series of estimated total population sizes $N_{total,t}$ against time for each iteration of the MCMC. This method results

in a posterior distribution of trends that translates the uncertainty in these abundance estimates to necessary uncertainty in population trend.

We assume that entry and detection are time-varying, while survival remains static throughout the study period. Given the longevity of false killer whales relative to the length of the study (Kasuya 1986) detectable changes in survival were not expected (see Bradford et al. 2018 for a discussion of this assumption).

Simulations of model performance

Simulations of this model's performance under varying conditions were conducted in Badger et al. (in review), and showed the model was robust to many realistic complications such as variable cluster-level space use and low detectability. Upon fitting the pseudospatial model to the full false killer whale dataset, we conducted further simulations to examine how the performance of this model is affected by some of the complications arising in our system. Following the procedure in Badger et al. (in review), we simulated animal movements, survey efforts, and resulting capture histories of 300 individuals comprising 3 social clusters. Starting locations for group-level movements and survey effort were chosen at random at each of 10 time steps, and subsequent movement and survey tracks were modeled as correlated random walks (using package `adehabitatLT` v. 0.3.26, Calenge 2006). For each of 10 time steps, surveys would detect individuals with detection probability of 0.2 if they were < 2 km from the survey vessel. The resulting capture histories were then fit to a pseudospatial JS model outlined above in a Bayesian framework. Six simulated individuals from each cluster were telemetered, and their locations used to estimate the cluster-level utilization distributions for which survey overlap was determined.

We considered three complications: unequal social cluster abundances, unequal telemetry tag deployment lengths, and population trends. Cluster abundances were either equal, with 100 individuals in each group, or unequal, distributed as 50, 100, and 150 individuals. Telemetry tag deployment lengths were either equal, at a mean of 62.86 days, or were pulled from a Normal distribution with parameters $\mu = 62.86$ days and $\sigma = 48.23$ truncated to > 0, the sample mean and standard deviation derived from available telemetry data. We also simulated positive and negative trends in abundance, with the slope of abundance over the time series as -5% or +5%.

This procedure was repeated 30 times per set of conditions (equal vs. unequal cluster abundances,

equal vs. unequal tag deployment lengths, positive vs. negative population trends) to observe the range of probabilistic outcomes in capture histories. For each model, we report the difference in posterior precision, the number of inaccurate abundance estimates (see below), and the proportion of iterations whose abundance estimates exhibit an inaccurate trend. We define an inaccurate abundance estimate as one where the 90% credible interval (CRI) of the posterior distribution did not contain the true population size. To detect inaccurate trends, we regressed derived abundance estimates over time for each iteration of the Bayesian Markov chain Monte Carlo (MCMC), and defined an inaccurate trend if the 90% credible interval of the resulting slope parameter distribution was distinct from 0.

Accounting for distinctiveness

Importantly, in using photo-identification for capture-recapture analyses, abundance estimates N_t will only reflect the number of distinctive individuals in the population, as nondistinctive individuals are excluded from our dataset (until they become distinctive). To estimate total abundance, we must adjust abundance estimates by the proportion of the population that is distinctive in each year (θ_t). Total abundance is then estimated by $N_{total,t} = N_t/\theta_t$, where $N_{total,t}$ is the abundance of all individuals (distinctive and nondistinctive) at each capture occasion t.

Unfortunately, there is no information in the resight data alone to inform these parameters. Therefore, we estimated θ_t using photographs taken from encounters by CRC where the number of nondistinctive and distinctive individuals were determined in each group to estimate the proportion of distinctive individuals in the population in each year, expanding on the approach of Bradford et al. (2018). We fit a binomial (logit link) generalized additive model (GAM) to these data with year as a covariate to get a smoothed function of proportion distinctive by year. We then extracted the values of the linear predictor, coefficients, and variance-covariance matrix to generate a prior on the time series of θ_t. At each iteration of the capture-recapture MCMC, we pulled predicted curves from this model (Figure 4) to appropriately propagate uncertainty in θ_t to the $N_{total,t}$ calculation in adjusting our abundance estimates to account for nondistinctive individuals in the population. While information on counts of nondistinctive and distinctive individuals were not available for the other contributors, CRC encounters comprise the majority of the dataset and thus were deemed reasonable to apply to all encounters.

Model fitting & selection

This model, as well as a null model without the overlap variable, was fit to the annual encounter histories compiled for each individual in the photo-identification data set. For further exploration of model formulations of a conventional Jolly-Seber model fit to a previous version of this dataset, see Bradford et al. (2018). A Bayesian framework was used for model fitting, selection, and inference using the software JAGS 4.2.0 through the `R` interface `rjags` (`R` version 4.2.2; Plummer 2003, R Core Team 2020, Plummer 2018). We used an informative Beta(8,2) prior for survival rate ϕ (given the longevity of false killer whales), but otherwise uninformative priors, namely Uniform(0,1) distributions, for probabilistic parameters constrained to [0,1], The parameter describing the effect of overlap, δ, was given a diffuse Normal(0,1000).

Markov chain Monte Carlo (MCMC) was used to sample the posterior distributions of the parameters of interest. For each model, we ran three chains with different sets of initial values. The first 10,000 MCMC samples were discarded as the burn-in after having checked that convergence was satisfactory. Convergence of chains to stationary distributions was visually evaluated using sample path plots in conjunction with the Brooks-Gelman-Rubin diagnostic $\hat{r}$ (Brooks and Gelman 1998), with values close to 1.00 indicating adequate convergence. Chains were then run for 20,000 iterations after burn-in, and every tenth iteration was retained for a total of 2,000 MCMC samples used for inference. We determined that a covariate had an effect if a 95% credible interval (CRI) of the posterior distribution of that parameter did not include 0. We assessed support for inclusion of overlap using a measure of out-of-sample predictive ability of each model, the Widely Applicable Information Criterion (WAIC, Watanabe 2010), where a model with a smaller WAIC is judged a better fit.

Results

Collectively, survey efforts resulted in a photo-identification dataset of 202 individuals comprised of high-quality photographs that can be translated into encounter histories for use in capture-recapture models (details of this process can be found in Bradford et al. (2018)). The telemetry dataset included 53 deployments (Cluster 1, n = 27; Cluster 2, n = 5; Cluster 3, n = 16; Cluster 4, n = 5), ranging from 12-199 days of data (mean = 62.5/median=48.8 days) that were analyzed for population-level space use.

We fit the pseudospatial and conventional Jolly Seber model to encounter histories of 202

individuals over 23 years, and found strong support for the pseudospatial model formulation (Δ WAIC = 33.2). The posterior distribution of the parameter describing the effect of overlap on detection probability was distinctly positive (Figure 7), indicating that years with higher survey overlap with cluster-level space use had higher detection probabilities.

We found a negative trend in abundance estimates with time (posterior mean slope = -3.51, Figure 5), though the 95% CRI of the posterior distribution of this derived parameter spanned 0 (95% CRI = [-8.40, 2.04], Figure 6). The posterior distribution indicates that there is an 81.4% probability the trend is negative over the entire time-series, though the relationship between abundance estimates and time does not appear linear (Figure 5). We also report the trend for the last 10 years of abundance estimates, which is distinctly negative (posterior mean = -5.53, 95% CRI = [-9.91, -1.61], Figure 6), with a 99.6% probability that the trend is negative over the 10 year period.

Simulation results

We found that performance of the pseudospatial model is only marginally compromised by variable tag durations, population trends, and cluster-level abundances (Table 1). When fit to data generated with variable population trends, we find that the pseudospatial model can better estimate population size and trend, whether the population is growing, shrinking, or remaining constant. While the pseudospatial model still performs well under variable cluster-level abundances, with 90% of fits including an accurate estimate of true abundance N, accurate trend detection was less frequent at only 69%. Further, upon post-processing calculation of cluster-specific abundances, we found that the model did not accurately estimate the number of individuals in each cluster. As the total population size was still accurate, this is likely linked to cluster "label-switching" within and among MCMC chains.

Cluster-level results From our simulations, it was clear that the full population model as formulated would not suffice for estimating true cluster-level abundances. So, we instead ran the model separately for each cluster. Many parameters were inestimable for Clusters 2 and 4 that have more sparse data, resulting in poor convergence of MCMC chains in Cluster 2. Clusters 1, 3, and 4, which make up the most encountered groups in the dataset, have similar estimated trends as the full population (Table 2). Cluster 2 is the only cluster with an estimated increasing trend, but the

estimated abundances in this model had poor convergence ($\hat{r}$ ¿ 1.1).

Discussion

We estimated abundance for the MHI insular population of false killer whales using a pseudospatial model that accounts for animal availability in detection and found the population has likely been in decline for at least the last 10 years (Figure 6), a result that is consistent among candidate models. Clusters 1 and 3, the most represented social clusters in the dataset, exhibited similar trends in abundance. These results constitute the first trend estimates for this population, which has been listed as endangered since 2012.

Simulated data sets, in this abundance analysis and in previous analysis of a shorter time series (Badger et al., in review), showed that using the pseudospatial method will yield higher accuracy and precision in estimating abundance and trend than conventional models, and its performance is relatively unaffected by varying population trends, poor characterization of cluster-level space use, low detectability, and unequal social cluster abundances and tag deployment durations. These simulations were designed to target specific sample size considerations in this dataset. First, there is a vast disparity in telemetry sample sizes among clusters, such that space use for Clusters 2 and 4 are informed by only 5 tagged false killer whale groups each (after accounting for pseudoreplication among tagged pairs), whereas Clusters 1 and 3 are informed by 27 and 16 tagged groups, respectively. Not only does this sparsity inhibit our ability to adequately characterize cluster-level space use for Cluster 2 and 4, but precluded us from estimating time-specific space use kernels and thus limited us to assume constant space use over time. Fortunately, simulations indicate that if this limitation mischaracterizes space use, it does not reduce the quality of the abundance estimates to less than that of the conventional CR model. Our model did estimate a positive effect of overlap on detection probabilities, indicating that our calculated overlap variable is at the very least not random with respect to observed encounter rates, and likely explains some variability in availability. Clusters 2 and 4 are also underrepresented in the visual sighting data, affecting our ability to estimate a cluster-specific trend. Because Clusters 2 and 4 are generally a small portion of the data regardless of sampling site, it is reasonable to assume that these social clusters have a smaller population size than Clusters 1 and 3. Our simulation exercise indicates that this discrepancy in cluster abundance will not greatly impact estimation of the full population abundance. Regardless, these do highlight

the need for targeted field efforts in the known ranges of Clusters 2 and 4 to better understand how cluster-level dynamics scale up to the population-level decline.

We note a decline in the proportion of distinctive individuals throughout the time series, from roughly 80% of the individuals in encountered groups designated as distinctive in early years to 70% in 2021 (Figure 4). We expect that this trend is an artifact of sampling limitations rather than a true decline. Advancements in camera technology have allowed field photographers to have greater group coverage and higher quality photos of all individuals. Capturing quality identification photographs of small, fast-moving and generally nondistinctive individuals increased later in the time series. If the proportion of distinctive individuals in the first portion of the time series is biased high as a result, this may account for the coinciding increasing trend in abundance for those years (1999-2002). It is also possible that there has been a shift in the age structure in the population, with an increase in younger cohorts (less distinctive) coincident with a decrease in older cohorts (more distinctive).

The estimated decline in abundance is consistent among candidate models (conventional CR model 95% CRI:[-11.574, 2.936]). Bradford et al. (2018) did not find a similar trend when fitting a conventional CR model to sightings data from 2000-2015. While some may argue this is an artifact of the modeling process and not reflective of a true decline in abundance, there is more evidence that the difference stems from the information in the additional data collected since 2015. Badger et al. (in review) fit the pseudospatial model framework to the same data as Bradford et al. (2018), the sightings data from 2000-2015 from one source (CRC), and also did not find a substantial trend in abundance estimates. Further, we fit the pseudospatial model to the full time series of data contributed by CRC, which is the largest portion of the dataset with the best-described effort information, from 1999-2021 and found a similar decline with that data partition.

This analysis could be improved in several ways for future abundance estimation. Firstly, the spatiotemporal nature of the data generating process could be more explicitly defined and incorporated into the modeling framework. Current development of a continuous-time spatially-explicit CR model, incorporating a time inhomogenous Markov-modulated Poisson process, could address this point (Johnson and Robinson, in prep). Additionally, as we were only able to estimate a static cluster-level space use layer for availability using the telemetry data, further information on temporal variation in animal availability to survey efforts would improve our method. For example,

passive acoustic monitoring data could be easily incorporated into this framework, though currently limited to provide presence-only information in small areas over long time periods or broader regions over shorter time scales. Incorporating additional auxiliary data into abundance estimation will greatly aid not only recovery metric accuracy and power to detect trends, but also stakeholder confidence in management actions for this population.

References

Andrews RD, Pitman RL, Ballance LT. 2008. Satellite tracking reveals distinct movement patterns for Type B and Type C killer whales in the southern Ross Sea, Antarctica. *Polar Biology* **31**: 1461–1468. ISSN 07224060.

Baird R, Hanson M, Schorr G, Webster D, McSweeney D, Gorgone A, Mahaffy S, Holzer D, Oleson E, Andrews R. 2012. Range and primary habitats of Hawaiian insular false killer whales: informing determination of critical habitat. *Endangered Species Research* **18**: 47–61. ISSN 1863-5407.

Baird RW, Gorgone AM, McSweeney DJ, Webster DL, Salden DR, Deakos MH, Ligon AD, Schorr GS, Barlow J, Mahaffy SD. 2008. False killer whales (Pseudorca crassidens) around the main Hawaiian Islands: Long-term site fidelity, inter-island movements, and association patterns. *Marine Mammal Science* **24**: 591–612. ISSN 08240469.

Bradford AL, Baird RW, Mahaffy SD, Gorgone AM, McSweeney DJ, Cullins T, Webster DL, Zerbini AN. 2018. Abundance estimates for management of endangered false killer whales in the main Hawaiian Islands. *Endangered Species Research* **36**: 297–313. ISSN 16134796.

Brooks SP, Gelman A. 1998. General methods for monitoring convergence of iterative simulations. *Journal of Computational and Graphical Statistics* **7**: 434–455. ISSN 15372715.

Calenge C. 2006. The package adehabitat for the R software: tool for the analysis of space and habitat use by animals. *Ecological Modelling* **197**: 1035.

Conn PB, Ver Hoef JM, McClintock BT, Johnson DS, Brost B. 2022. A GLMM approach for combining multiple relative abundance surfaces. *Methods in Ecology and Evolution* **13**: 2236–2247. ISSN 2041210X.

Douglas DC, Weinzierl R, C Davidson S, Kays R, Wikelski M, Bohrer G. 2012. Moderating Argos location errors in animal tracking data. *Methods in Ecology and Evolution* **3**: 999–1007. ISSN 2041210X.

Dujon AM, Lindstrom RT, Hays GC. 2014. The accuracy of Fastloc-GPS locations and implications for animal tracking. *Methods in Ecology and Evolution* **5**: 1162–1169. ISSN 2041210X.

Fleming CH, Fagan WF, Mueller T, Olson KA, Leimgruber P, Calabrese JM. 2015. Rigorous home range estiamtion with movement data: a new autocorrelated kernel density estimator. *Ecology* **96**: 1182–1188. ISSN 0012-9658.

Gelman A, Carlin JB, Dunson DB, Vehtari A, Rubin DB. 2013. *Bayesian Data Analysis*. Boca Raton, FL: Chapman and Hall/CRC, 3rd edition.

Johnson DS, London JM, Lea MA, Durban J. 2008. Continuous-time correlated random walk model for animal telemetry data. *Ecology* **89**: 1208–1215.

Jolly GM. 1965. Explicit estimates from capture-recapture data with both death and immigration-stochastic model. *Biometrika* **52**: 225. ISSN 00063444.

Kasuya T. 1986. Fase killer whales. In *Report of investigation in search of solution for dolphin-fishery conflict in the Iki Island Areas*. Tokoyo, Japan: Japan Fisheries Agency.

Kranstauber B, Cameron A, Weinzerl R, Fountain T, Tilak S, Wikelski M, Kays R. 2011. The Movebank data model for animal tracking. *Environmental Modelling and Software* **26**: 834–835. ISSN 13648152.

London JM. 2020. pathroutr: An R Package for (Re-)Routing Paths Around Barriers.

Plummer M. 2003. JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling.

Plummer M. 2018. rjags: Bayesian graphical models using MCMC. R package version 4-8.

R Core Team. 2020. R: A language and environment for statistical computing.
URL `https://www.r-project.org/`

Schorr GS, Baird RW, Bradley Hanson M, Webster DL, McSweeney DJ, Andrews RD. 2010. Movements of satellite-tagged Blainville's beaked whales off the island of Hawai'i. *Endangered Species Research* **10**: 203–213. ISSN 18635407.

Seber GAF. 1965. A note on the multiple-recapture census. *Biometrika* **52**: 249–260.

Vehtari A, Gelman A, Simpson D, Carpenter B, Burkner PC. 2021. Rank-Normalization, Folding, and Localization: An Improved (Formula presented) for Assessing Convergence of MCMC (with Discussion)*†. *Bayesian Analysis* **16**: 667–718. ISSN 19316690.

Watanabe S. 2010. Asymptotic equivalence of Bayes cross validation and Widely Applicable Information Criterion in singular learning theory. *Journal of Machine Learning Research* **11**: 3571–3594.
URL `http://arxiv.org/abs/1004.2316`

Yano KM, Oleson EM, Mccullough JLK, Ballance LT, Hill MC, Bradford AL, Allen AN, Joyce TW, Moore JE, Henry AE. 2018. Cetacean and Seabird Data Collected During the Winter Hawaiian Islands Cetacean and Ecosystem Assessment Survey (Winter HICEAS), January-March 2020. *NOAA Technical Memorandum NMFS-PIFSC-72* .
URL `https://doi.org/10.25923/ehfg-dp78`

Tables

Table 1: Performance of the pseudospatial and conventional Jolly-Seber open population model fit to 30 simulated datasets under varying conditions. % with true N and % with true trend refer to the accuracy (90% CRI) of posterior distributions of abundance and the slope parameter in a regression fit to yearly abundance estimates. PS and C denote that the model fit was the *psuedospatial* or *conventional* model, respectively. The change in posterior precision was measured as the ratio of posterior standard deviations between two competing models.

Simulated condition	Model	% true N	% true trend	Change in posterior precision
Variable tag duration	PS	87%	71%	-
Constant tag duration	PS	90%	74%	1.01
Decreasing trend	C	82%	58%	-
Decreasing trend	PS	85%	62%	1.18
Increasing trend	C	80%	58%	-
Increasing trend	PS	86%	64%	1.19
Variable cluster abundances	PS	90%	69%	-
Constant cluster abundances	PS	92%	75%	1.03

Table 2: Cluster-level trends estimated using the pseudospatial Jolly Seber.

Cluster	Posterior mean trend	95% CRI	Probability of decline over time series
1	-2.34	[-7.561, 1.09]	0.911
2	1.951	[-0.265, 3.87]	0.038
3	-0.352	[-1.393, 0.533]	0.655
4	-0.938	[-2.13, 0.225]	0.946

Figures

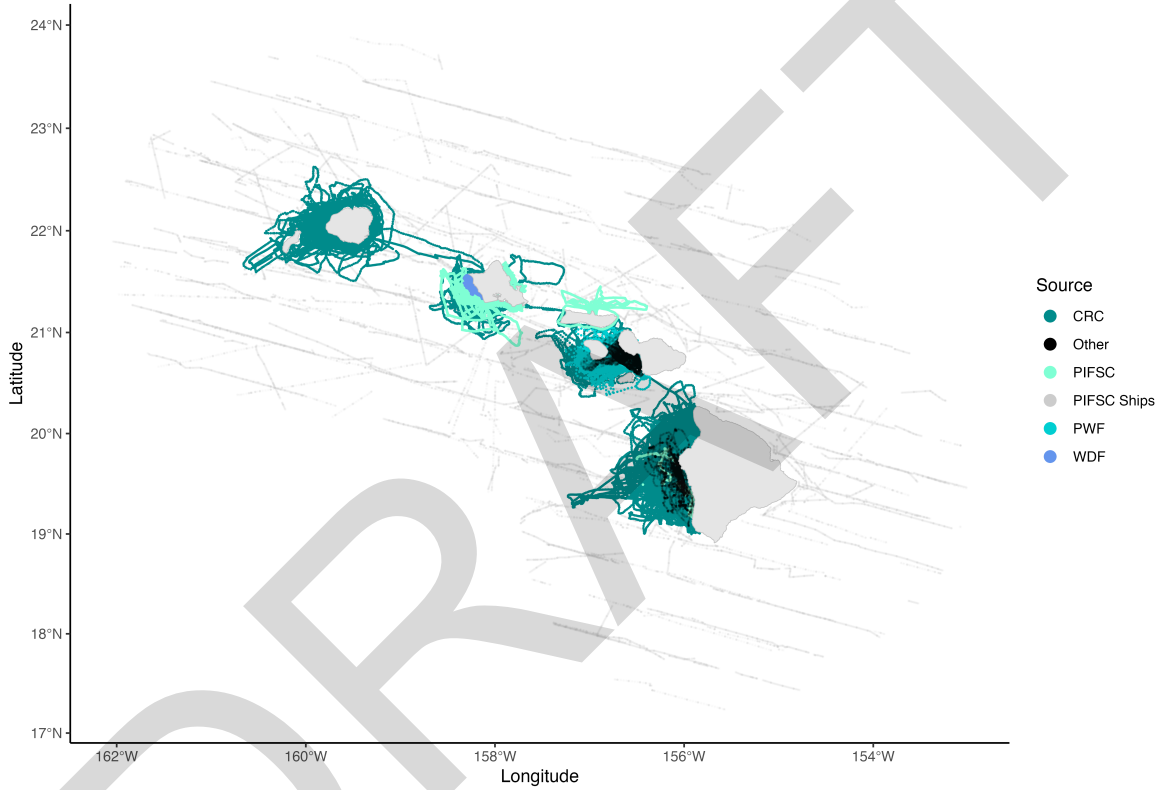

Figure 1: Tracks from surveys conducted by Cascadia Research Collective, Pacific Islands Fisheries Science Center (systematic ship surveys in light grey), Pacific Whale Foundation, and Wild Dolphin Foundation, as well as other opportunistic contributors, from 1999-2021 in the Main Hawaiian Islands.

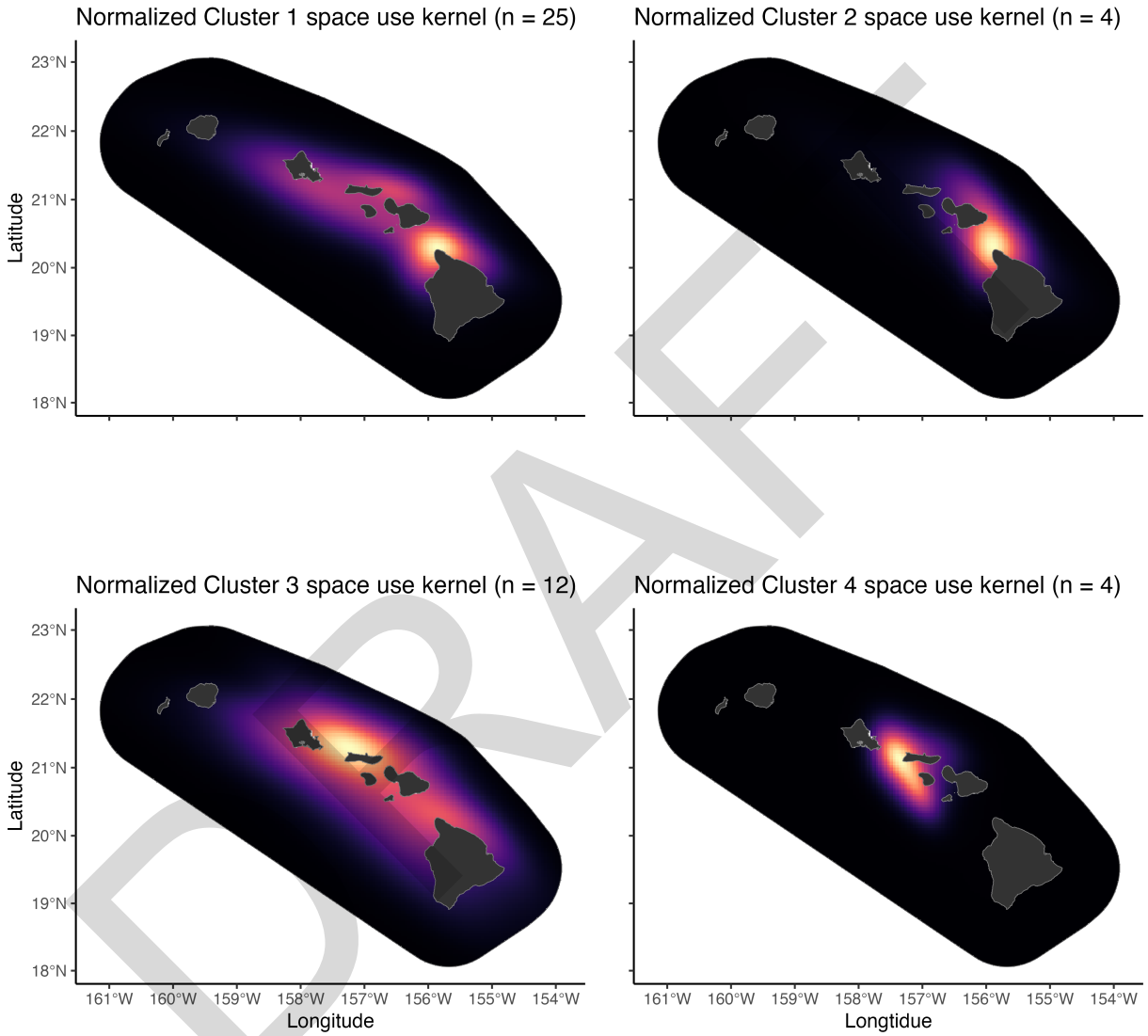

Figure 2: Cluster-level space use of Main Hawaiian Island insular false killer whales determined from kernel density estimators of location data from 52 tags deployed on individuals from 4 social clusters from 2007-2022.

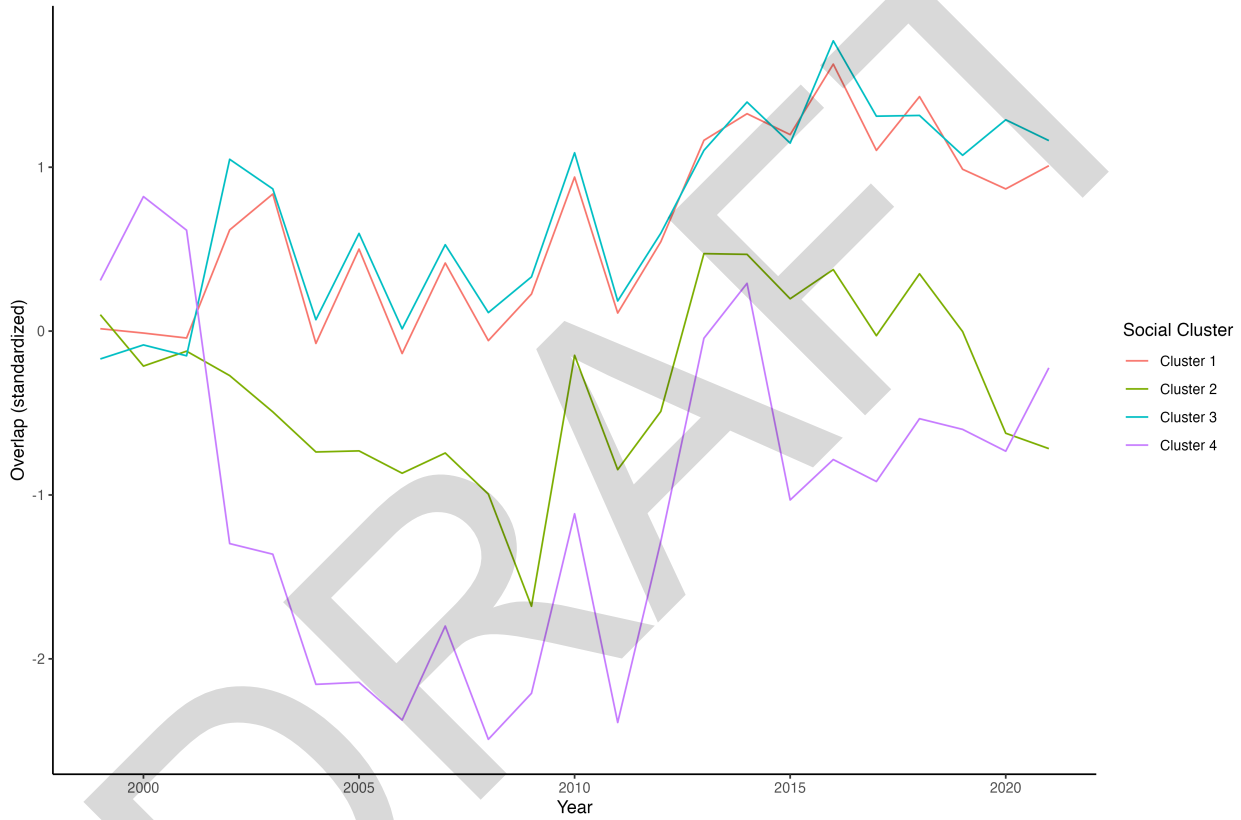

Figure 3: Overlap of survey efforts and social cluster-level space use of the Main Hawaiian Island insular population of false killer whales from 1999-2021.

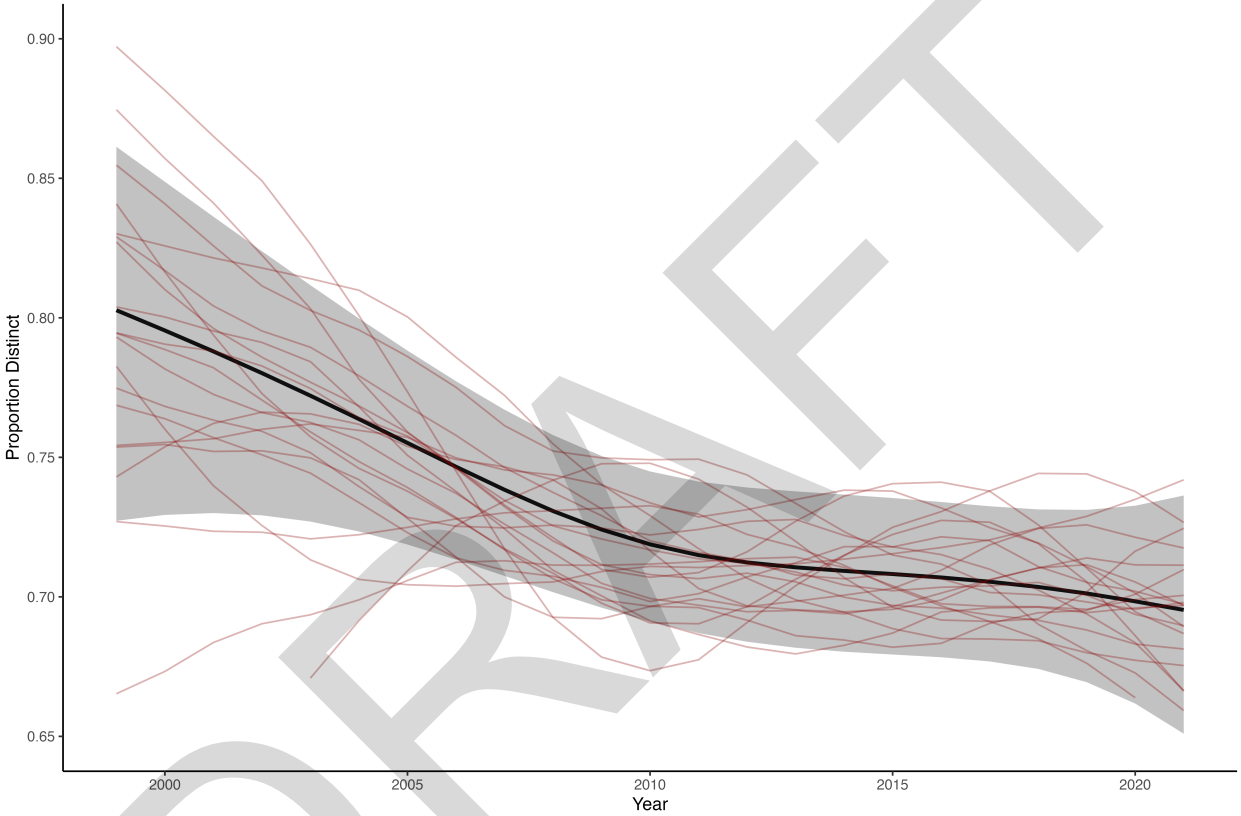

Figure 4: Output of generalized additive model fit to the proportion of distinctive individuals in sightings by the Cascadia Research Collective from 1999-2021. Black line and grey shaded area is the mean and 95% confidence interval of predicted relationship between proportion distinctive and year. Red lines are sample draws of the multivariate normal distribution described by model output, the prior distribution used for proportion distinctive in the capture-recapture MCMC.

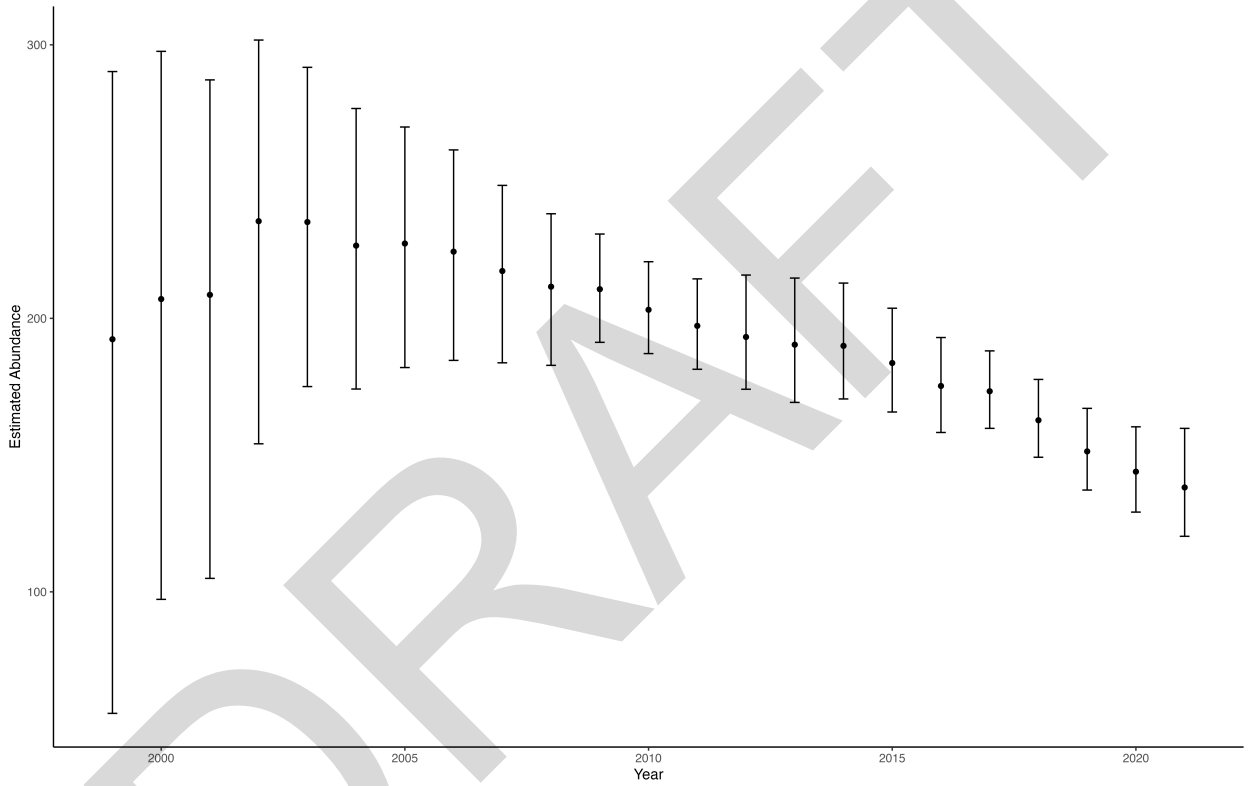

Figure 5: Abundance of Main Hawaiian Islands insular false killer whales, 1999-2021, estimated using a pseudospatial Jolly-Seber model. Error bars indicate 95% credible intervals.

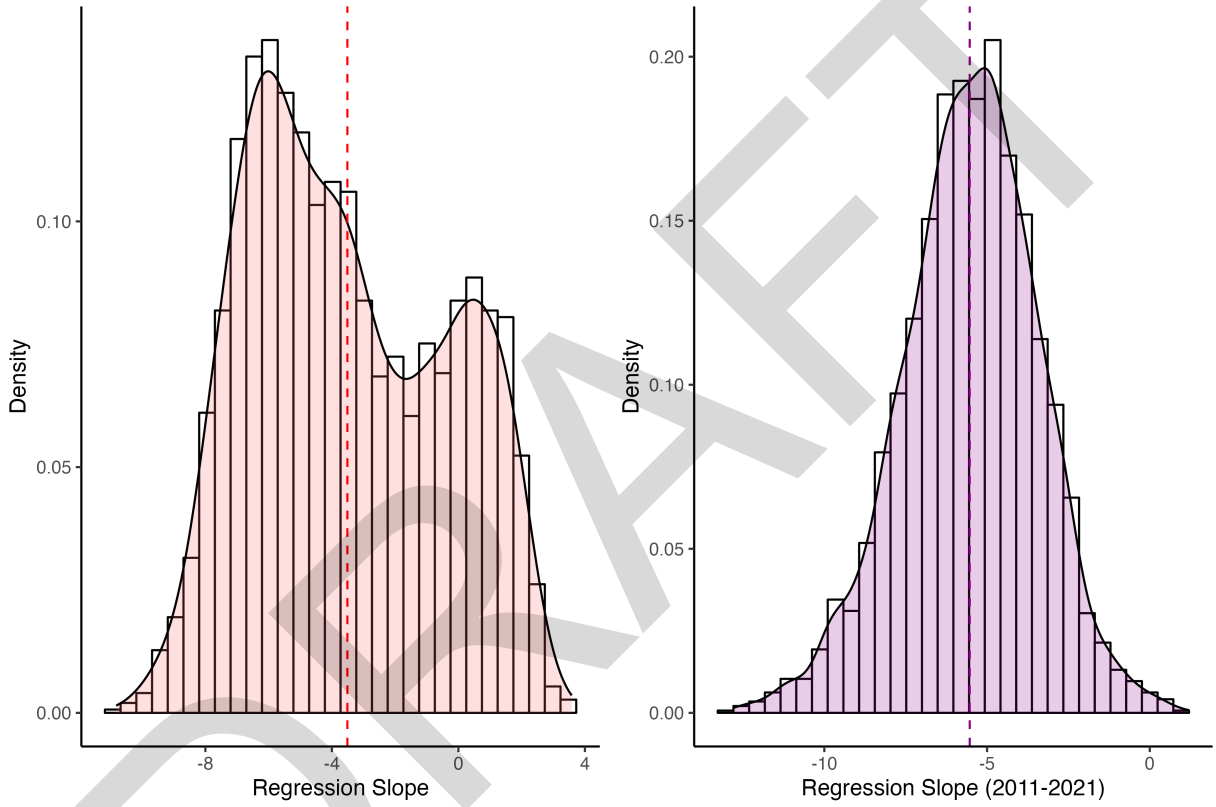

Figure 6: Posterior distributions of estimated population trends for the entire Main Hawaiian Islands insular population from 1999-2021 (red, left) and 2011-2021 (purple, right). Dashed lines indicate posterior mean.

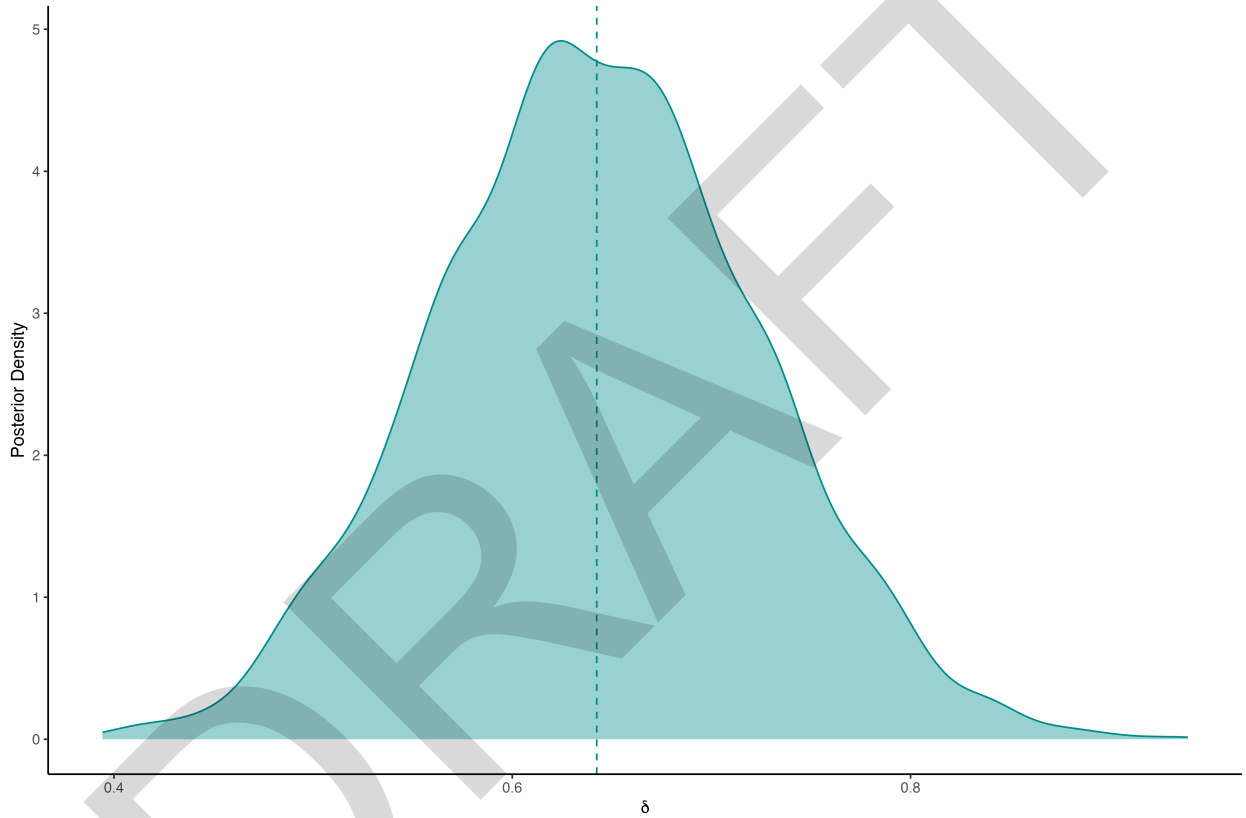

Figure 7: Posterior distributions of the parameter describing the effect of the overlap between survey effort and animal space use on detection probability. Dashed line indicates posterior mean.