FINAL REPORT
A power analysis and recommended study design to directly detect population-level consequences of acoustic disturbance

Jeffrey Moore and Jay Barlow
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
phone: (858) 546-7161 fax: (858) 546-7003 email: jeff.e.moore@noaa.gov
phone: (858)546-7178 fax: (858)546-7003 email: jay.barlow@noaa.gov

Erin Falcone and Greg Schorr
Marine Ecology and Telemetry Research
2420 Nellita Road NW
Seabeck, WA 98380
phone: (206) 550-9806 email: efalcone@marecotel.org

David Moretti
Naval Undersea Warfare Center
Newport, RI 02841
phone: (360) 943-7325 ext. 101 email: david.moretti@navy.mil

K. Alexandra Curtis
Ocean Associates, Inc.
Under contract to Southwest Fisheries Science Center
National Marine Fisheries Service
National Oceanic and Atmospheric Administration
8901 La Jolla Shores Drive
La Jolla, CA 92037
phone: +1-858-546-5625 email: alex.curtis@noaa.gov

Award Number: N0001415IP00088 / N0001415WX01764 / N000141512899)
https://swfsc.noaa.gov/textblock.aspx?Division=PRD&id=1222&ParentMenuId=148

LONG-TERM GOALS

The overall project goal was to provide advice to the Navy on monitoring health and status of a beaked whale population in a high-use area for Navy exercises employing mid-frequency active sonar. Recommendations address the following question: "How do we quantify potential human impacts on the health of beaked whale populations?"

OBJECTIVES

Our project objectives were to assess the population of Cuvier's beaked whales, *Ziphius cavirostris*, that use the Southern California Anti-submarine Warfare Range (SOAR; Fig. 1), and to apply the results in a power analysis to inform sampling design for continued Navy monitoring of beaked whale populations. We specifically aimed to (1) estimate abundance and trend for the population of *Ziphius* that use the San Nicolas Basin, based on photo-identification data; (2) conduct a statistical power analysis based on those results to determine the power of ongoing monitoring (photo-identification and passive acoustic monitoring programs) to detect trends in population health and status of beaked whales in the area; and (3) make recommendations for more effective continued monitoring to detect trends in population health and status.

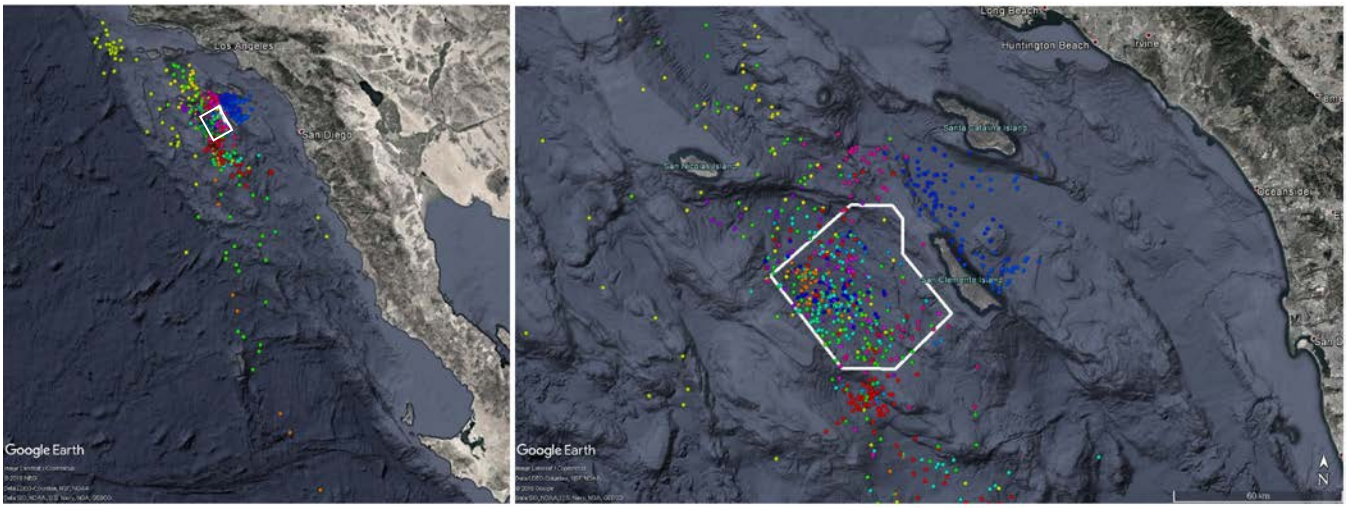

Figure 1. Daily positions of Ziphius cavirostris satellite-tagged within SOAR (outlined in white). Left panel shows all recorded locations; right panel shows positions within the Southern California Bight (right panel).

APPROACH

We proposed to conduct a power analysis that would inform sampling design for continued Navy monitoring of beaked whale populations on SOAR, based on results of population models fit to photo-identification and acoustic monitoring data for the population of *Ziphius*.

Cetacean monitoring data have been collected in the vicinity of the Navy's Southern California Offshore Range (SCORE) from multiple survey programs of the past several to many years, with intensive monitoring on SOAR (in the San Nicolas Basin) via two synergistic programs: photo-identification and passive acoustics. Passive acoustic monitoring (Moretti et al.) has been conducted using a fixed array of bottom-mounted hydrophones since 2006, and with beaked-whale-specific classifiers since 2010. The monitoring system is part of the Naval Undersea Warfare Center's (NUWC) Marine Mammal Monitoring on Ranges (M3R) program that leverages the SOAR hydrophone array. Cascadia Research Collective and now Marine Ecology and Telemetry Research (PIs: Falcone and Schorr) has collected photo-identification (mark-recapture) data for *Ziphius* in SCORE since 2006, with a focus on SOAR, where a partnership with the M3R program supports use of the hydrophone array to enhance the photo-ID team's ability to detect and locate groups of *Ziphius*. Adult *Ziphius* generally have distinctive markings and are thus well-suited to photo-identification

techniques, provided an adequate sample of images can be obtained over time. Photo-ID data for *Ziphius* include several years of pilot study (2006-2009) followed by directed sampling since 2010. Over the course of the study, surveys have been conducted during ten different months of the year, though effort since 2010 has been disproportionately focused in winter and spring months to balance previous efforts (2006 – 2009) in the region, which occurred predominantly in the late summer and fall. Surveys are conducted using a 6m rigid-hulled inflatable boat (RHIB) in partnership with the M3R program. Photos are scored for distinctiveness of markings, and weather/sea conditions during effort are recorded continuously. Details of the field methodology have been described by Falcone et al. (2009) and in other reports (e.g., Falcone et al. 2011).

Ultimately, we focused primarily on analysis of the photo-ID data for two reasons. First, as data were explored during this project, the project team gained an increasing appreciation for the fact that the acoustic and photo-ID data ultimately measure different, albeit complementary, phenomena. The acoustic array data provide high-resolution information (spatially and temporally) about what is occurring on the SOAR range at a particular point in time. The photo-ID data provide information about individuals from the broader population that occasionally uses SOAR. As such, abundance estimates from the two data sets are not comparable (they measure different things). This may or may not also be true of trend estimates; thus, we cannot assume that long-term use patterns on SOAR are the same as long-term trends in abundance for the broader population (e.g., Whitehead and Gero, 2015) although if this were determined to be the case, *then* the acoustics data could be used to help improve some population parameters estimated by the photo ID data. Moreover, given the amount of data collected from the acoustic array, the project team ultimately felt that a power analysis of the acoustic data was not necessary, because the data effectively represent a census (rather than a sample) of beaked whale use on SOAR such that statistical inference is not needed. The following therefore applies to the photo-ID data.

The implementation of the project can be broken down into several tasks, as summarized below:

- Task 1 involved discussion and exchange of data and associated knowledge (Moore, Barlow, Falcone, Schorr, Moretti, and contractor K. Alexandra Curtis), data exploration (Curtis, Falcone, Schorr, Moretti, and Moore), and quality control (Curtis and Falcone).
- Task 2 involved identifying and fitting a set of candidate open-population mark-recapture models for estimating abundance and trend information from the photo-ID data (Curtis and Moore).
- Task 3 involved designing and conducting a power analysis to assess and make recommendations for improving power to detect trends in population status (Curtis and Moore). Both frequentist and Bayesian approaches were considered for the power analysis.
- Task 4 involved project coordination. Participants in project meetings varied according to specific needs, but included Moore, Barlow, Falcone, Schorr, Moretti, and Curtis.

WORK COMPLETED

The sparse nature of the mark-recapture (photo-ID) data collected over the past decade in the San Nicolas Basin presented considerable challenges for model fitting and hence population and trend estimation. Nevertheless, we were able to generate preliminary mark-recapture-based estimates of apparent annual survival (including death and emigration), abundance, and population growth rate for the population of *Ziphius* that use the San Nicolas Basin on an ongoing basis. We used the resulting

estimates to conduct a power analysis to provide insights into data requirements and model choice to support robust inference about population trends and demographic parameters such as survival. We detail the work completed to achieve each of these key outcomes below.

Estimation of abundance and demographic rates
Our approach to estimating population size and demographic rates from the photo-ID data included three steps:

- Data exploration and quality control;
- Model development using Cormack-Jolly-Seber (CJS) models in a Bayesian framework, with emphasis on modeling capture probability (p; CJS results were also used to estimate apparent annual survival, henceforth just "survival" or φ); and
- Model development of Jolly-Seber models in a Bayesian framework, using the term for p resulting from model exploration in the CJS framework, to estimate abundance and population growth rate (λ).

Data quality control led to reprocessing of effort data and corrections to the photo-ID database. Scripts written for effort processing and data merging and selection for population modeling will greatly expedite future updates and continuations of modeling efforts based on the photo-ID data.

We designed criteria for including photo-ID records in the capture histories for model fitting to minimize the chance of a single animal being recorded as two different individuals, and to constitute a relatively homogeneous population in terms of identifiability and survival. We selected records of non-calf individuals with minimum criteria for photo quality and distinctiveness of markings. Capture data were aggregated into nine annual occasions from 2007 to 2015, with each annual occasion beginning August 1 to maximize temporal separation between occasions. We analyzed left- and right-side capture histories separately. Discovery curves and capture frequencies are provided Figure 2 and Table 1.

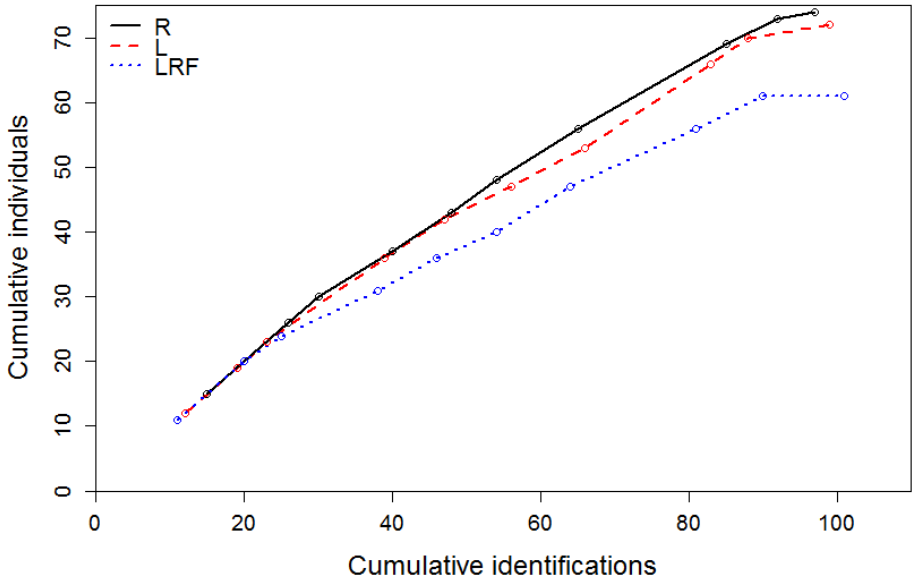

Figure 2. Discovery curves of total individuals versus number of identifications made each year from 2007 to 2016 (including partial occasion from August 1, 2016 to December 31, 2016).

Table 1. Capture frequencies of Ziphius cavirostris, total number of individuals (n), and average captures per individual (x) for each annual capture history used in our mark-recapture analysis The LRF capture history comprises individuals that were photographed from both sides within a sighting or had fin damage that would be recognizable from either side.

Capture history	n	x	1	2	3	4	5
Right-sided	73	1.26	55	15	4	–	–
Left-sided	70	1.26	57	10	2	–	1
LRF	61		39	17	4	–	1

We focused model development on characterizing capture probability *p*, because this parameter is subject to substantial bias as a result of temporal and individual heterogeneity, strongly influences population size estimates, and has varied considerably over the mark-recapture time series due to interannual variability in hours of effort completed, weather conditions during sampling, and timing of sampling in the year (Fig. 3). The latter may be important because, based on acoustic detection records, *Ziphius* is suspected to exhibit a marked decrease in use of the study area in summer and early fall (Baumann-Pickering et al. 2015). Exploratory data analysis also indicated potential individual heterogeneity in *p* (see also Falcone and Schorr 2014), which can lead to substantial underestimation of population size if ignored. Survival, which includes death and emigration, was modeled as constant among years and individuals.

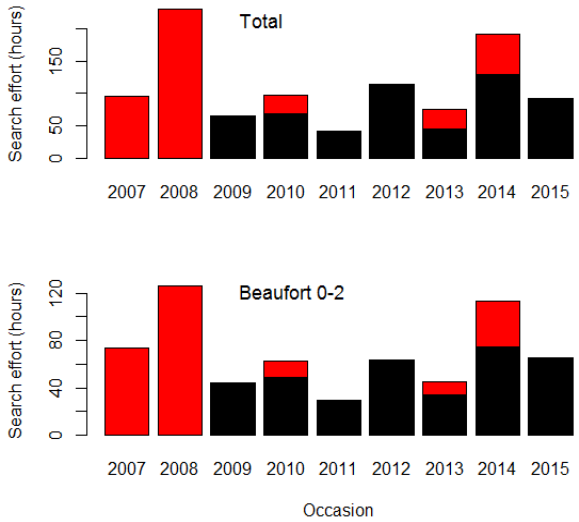

Figure 3. Hours of search effort for Ziphius cavirostris by season per occasion, for total search effort (top panel) and search effort at Beaufort Sea States 0 to 2. Red represents "summer" effort, defined as effort conducted during July through October (based on relative acoustic activity of Ziphius on SOAR measured by the hydrophone array). Black represents non-summer effort (all other months). Annual occasions begin in August of the occasion year.

We first used MARK software (White and Burnham 1999), integrated into the R work flow via RMark (Laake 2013), to obtain maximum likelihood estimates of abundance, population growth rate, and survival. However, even simple models (without covariates) did not consistently converge for different data subsets (e.g., left- vs right-sided histories). Additionally, both RMark and MARK software proved insufficiently flexible for modeling capture probability (no individual heterogeneity, restrictive covariate model choice). Consequently, all further models were fitted in a Bayesian framework, using JAGS through the rjags and jagsUI packages in R (Kellner 2016, Plummer 2016, R Core Team 2015).

Due to the relative speed of convergence for CJS models compared to Jolly-Seber models for the *Ziphius* mark-recapture data (minutes to tens of minutes versus days to weeks), we used CJS models (based on Kéry and Schaub 2011, henceforth the BPA CJS model formulation) for development of capture probability models. In addition, since CJS models condition on first capture (such that initial capture and recapture probabilities are allowed to differ, because only recapture probability is modeled), we were able to base model development for p on a relatively enriched data set consisting of individuals that were either photographed from both sides within at least one sighting or had fin damage that would be recognizeable from either side (henceforth the LRF dataset). These animals had a higher recapture probability than the rest of the population (Figure 2 and Table 1), leading to improved model estimates of p in spite of the lower sample size. We used a model with temporal and individual random effects in p as the base model for two reasons. First, we knew temporal variation in search effort and season existed, and data exploration suggested individual heterogeneity may be a factor. Second, simulations by White and Cooch (2017) show that the bias in p (and thus abundance estimates in Jolly-Seber models) from including temporal and random effects when none exist is low in our parameter space (low p, close to ten occasions) and further decreases with increasing data collection, while bias from not modeling random effects where variation does exist does not change with increasing data collection. We then evaluated influence of search effort with and without season (as an interaction) on temporal heterogeneity, and influence of distinctiveness on individual heterogeneity. We assessed covariate signal (or lack thereof) by comparing estimates of variance for temporal (for season or effort) or individual (for heterogeneity) random effects in models that included covariates to those that omitted them. We also explored model performance for simulated data sets based on parameters similar to those estimated for the true data, and finally checked model performance for the single-sided capture histories that are the basis for population estimation using the Jolly-Seber model framework. The final chosen model for p included temporal and individual random effects and a standardized effort covariate (hours on the water in San Nicolas Basin, with depth greater than 800 m and at Beaufort Sea State 0 to 2 conditions).

We additionally explored whether Bayesian latent multinomial models that included both left- and right-sided capture histories, with sides linked where identity was confirmed as the same (e.g., photographed from both sides in one day), improved precision of estimates (Link et al. 2010, Bonner and Holmberg 2013, McClintock et al. 2013), since a high percentage of animals were photographed near-simultaneously from both sides at some point. Specifically, we compared results from CJS models that included temporal and individual random effects using (a) single-sided capture histories and the BPA CJS model formulation from the preceding paragraph, versus (b) capture histories from both sides with the latent multinomial model in the multimark package (McClintock 2015). The latter uses similar priors and parameterization within sides as the BPA formulation. We eliminated multimark as an option for further exploration using Jolly-Seber models, because although the multimark CJS results had narrower credible intervals for φ than the single-sided BPA CJS results did, the credible interval for p was wider from the multimark model than from the single-sided BPA model.

Finally, we incorporated the chosen model for p into a Jolly-Seber model framework that estimates p, φ, superpopulation (total number of individuals that were in the population over the full time series, including new entrants and those that died or emigrated during that time), and annual entry probability to the population (by immigration or birth) from the superpopulation. We continued to model φ as constant. Entry probability was modeled as constant after the first occasion. We used a Bayesian Jolly-Seber model framework that applies parameter-expanded data augmentation to estimate superpopulation (Royle and Dorazio 2008, Kéry and Schaub 2011). Population size at each time point, and thus also population growth rate, are derived parameters. Jolly-Seber models were fit to both left-

and right-side capture histories. Fitting superpopulation models proved challenging due to very slow mixing of the hyperparameters for probability of entry into the study area population (birth and immigration). This challenge in achieving convergence can be attributed to the sparseness of the data due to low *p*, and the nearness of most of the parameter estimates to an upper or lower bound: Markov Chain Monte Carlo (MCMC) samples for φ have a mode near one, and *p* is low, as are entry probabilities after the first occasion. We used data cloning with CJS and POPAN models in MARK through RMark to ensure that the parameters near boundaries (φ, *p*, and probability of entry after the first occasion) are identifiable (Lele et al. 2007, Lele et al. 2010, Cooch and White 2016). The model based on left-side capture histories converged more slowly still than that based on the right side due to an outlier individual in the left-side data set that was captured six times, inflating individual heterogeneity and thus abundance estimates and data augmentation requirements. During the CJS model step, we had also seen some evidence that models based on the left-side capture histories overestimate individual random effects, therefore underestimating *p* (which would overestimate abundance in the Jolly-Seber model framework). We focus on results from right-side capture histories. We were able to achieve convergence of parameter estimates (psrf values of less than 1.05) by using the autojags function in the jagsUI package, with three MCMC chains, an adaptation phase of 50,000 samples, 600,000 burn-in samples, and 400,000 iterations thinned to every 100th sample. This resulted in 12,000 samples, though effective sample size was much lower for most parameters due to autocorrelation.

Power analysis

Our experience with fitting open population models led us to choose a frequentist approach to power analysis. We focused the power analysis on characterizing power to detect trends in abundance, using maximum likelihood estimation of Pradel-lambda models in MARK with RMark. The Pradel-lambda model framework builds on the CJS model framework by using time reversal to estimate recruitment (or in this particular parameterization, population growth rate) in the same manner that φ is estimated from the forward moving set of capture histories in a CJS model. Fitting a simple Pradel-lambda model results in estimates of φ, *p*, and population growth rate.

We considered 21 scenarios in all, and simulated 1000 capture histories for each scenario. The first six scenarios included capture histories for slow and fast population decline (λ of 0.97 and 0.93 respectively, simulated as constant over time) monitored for 9 years at current effort levels (Scenarios 1 and 2), 18 years at current effort levels (Scenarios 3 and 4), or 9 years at double the annual effort (Scenarios 5 and 6). Survival (φ = 0.93 for slow decline and 0.9 for fast), *p* (0.063 for current effort and 0.12 for doubled), starting population size (130), temporal fixed (i.e., effort) and random effects ($\sigma_{tFE}^2 = \sigma_{tRE}^2 = 0.125$), and individual random effects ($\sigma_{iRE}^2 = 0.25$) were based on estimates arrived at in the preceding analysis. Scenarios 7 to 12 controlled for the influence of individual random effects by setting σ_{iRE}^2 to zero. Scenarios 13 to 18 controlled for influence of fixed vs temporal random effects by setting σ_{tFE}^2 to zero and σ_{tRE}^2 to 0.25. Scenarios 19 to 21 addressed intra-annual sampling design, characterizing the influence of temporal spacing of sampling within year and seasonality of *p* on power and confidence interval coverage. To simulate these scenarios, search effort was simulated at the monthly level and capture histories then aggregated to annual occasions for model fitting. Scenario 19 is based on slow population decline and nine years of data collection, with mark-recapture effort occurring every 4 months, and *p* and fixed and random effect variances adjusted such that, when aggregated to annual resolution, the capture histories have similar characeristics to those in Scenario 1. In Scenario 20, sampling occurred in three consecutive months each year. In Scenario 21, sampling occurred every four months, but with a higher temporal random effects variance and lower mean *p* to simulate seasonal effects on *p*.

Pradel-lambda models were fit to the resulting capture histories for each scenario with an effort covariate for p when $\sigma_{tFE}^2 > 0$, and with p modeled as constant when $\sigma_{tFE}^2 = 0$. For each simulation, we tracked whether (1) the model had converged, (2) a population decline had been detected (i.e., upper 95% confidence limit for λ was less than one), (3) true φ fell within the estimated 95% confidence interval, and (4) true λ fell within the estimated 95% confidence interval. Power and confidence interval coverage were then evaluated for each scenario based on those simulations for which the model converged. Ideally, we would have fitted Pradel-lambda models with temporal random effects to the simulated capture histories to more closely approximate our ability to capture temporal variability in the Bayesian models, but this added a layer of complication that we lacked the time to address.

RESULTS

Abundance and demographic rates

Our best estimate of φ for non-calf *Ziphius* using SOAR is drawn from the Bayesian CJS model (BPA formulation) for right-side capture histories, where p was modeled with an effort covariate and individual and temporal random effects. This estimate is preferable to the similar but less reliable estimate from the Jolly-Seber model, because the CJS model had much less trouble converging and is thus more stable. The mean estimate is 0.947 (sd=0.042), the 50^{th} percentile is 0.957 (interquartile range of 0.924 to .981), and the mode is 0.990 (50% HPDI is .955 to 0.998).

Our current best estimates for population abundance and population growth rates for non-calf *Ziphius* using SOAR are drawn from the Jolly-Seber model based on right-side capture histories, and are summarized in Figures 4 and 5 and Table 2. Given the strong right-skew of posterior distributions for annual abundance, it is not straighforward to present a point estimate for the population size. The posterior mode (most analagous to a frequentist maximum likelihood estimate) ranged from 144 to 174 across years; the posterior median ranged from 170 to 195; and the posterior mean ranged from 203 to 232. Given this and the wide credible intervals, a robust statement is that the population probably numbers in the low hundreds. The estimates for abundance and new entrants (N_t, N_{super}, and B_t) from the model have been multiplied by 1.12 to account for the percentage of successfully photographed animals per day of sampling that did not meet distinctiveness criteria (negligible uncertainty relative to that of abundance estimates). The updated abundance estimates generally agree with earlier estimates from closed population approaches (Falcone and Schorr 2014).

The estimate for λ suggests the population is stable to increasing slowly (mode of 2% annual increase). However, uncertainty is large and overlaps substantially with negative population growth rates as well. Additionally, we caution that due to the shift in effort early in the time series from predominantly summer effort to predominantly non-summer effort, combined with the failure of our models to detect the known seasonal variation in acoustic activity in the area (and thus most likely also in p), it is possible that a downward trend in the population has been masked by an unmodeled upward shift in p. We note that these estimates remain preliminary and will be updated with an additional year of data this fall prior to publication, which is expected to considerably improve precision and – with better convergence – accuracy of estimates.

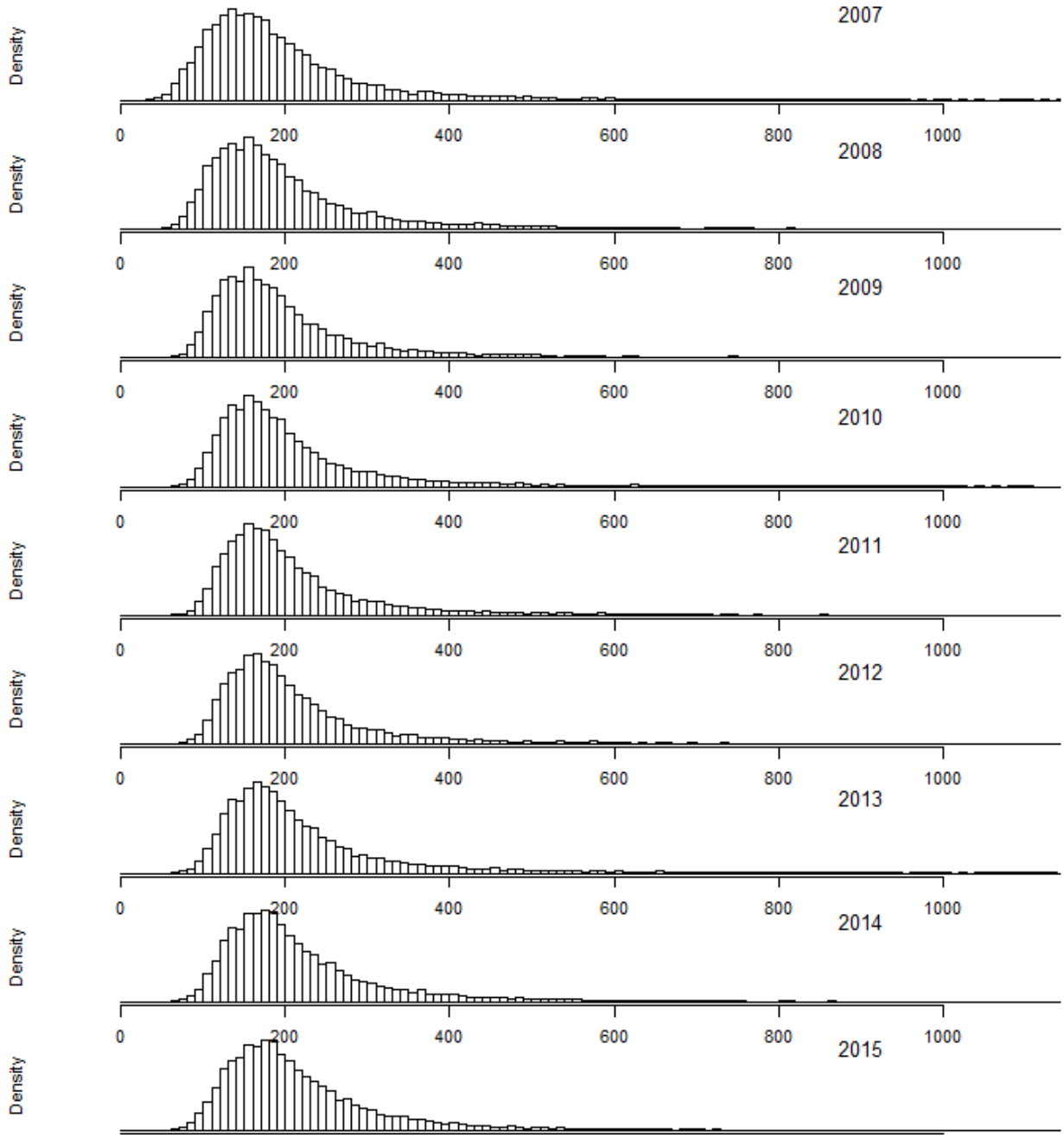

Figure 4. Population estimates for Ziphius using SOAR on an ongoing basis, from Bayesian Jolly-Seber model for right-side capture histories, where capture probability was modeled with temporal and individual random effects and an effort covariate. Abundance estimates were multiplied by 1.12 to account for non-calf Ziphius that did not meet distinctiveness criteria.

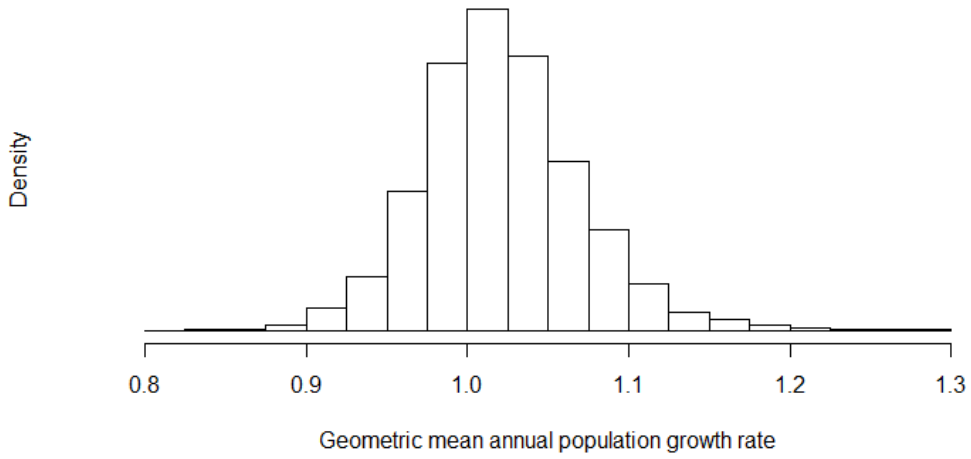

Figure 5. Geometric mean annual population growth rate (λ) estimates for Ziphius using SOAR on an ongoing basis, from Bayesian Jolly-Seber model for right-side capture histories, where capture probability was modeled with temporal and individual random effects and an effort covariate.

Table 2. Results for Bayesian Jolly-Seber model for right side capture histories, where capture probability was modeled with temporal and individual random effects and an effort covariate. φ is apparent annual survival, p is capture probability, a_{effort} is the effort coefficient for p, N_{super} is superpopulation abundance, λ is the geometric mean annual population growth rate, N_t is annual abundance, $B_{2:t}$ is number of new entrants to the population per year (from the second year onward). Subscript t=1 corresponds to 2007-08 (see text). σ_{tRE}^2 and σ_{iRE}^2 are variances on the logit scale for temporal and individual random effects in p. N_t, N_{super}, and B_t estimates from the model were multiplied by 1.12 to account for non-calf Ziphius that did not meet distinctiveness criteria. $H_{x,L}$ and $H_{x,U}$ are the lower and upper bounds of the x% highest probability density interval, P_x is the x^{th} percentile of the posterior distribution for each parameter, and SD is standard deviation.

	$H_{90,L}$	$H_{50,L}$	Mode	$H_{50,U}$	$H_{90,U}$	P_5	P_{25}	P_{50}	P_{75}	P_{95}	Mean	SD
φ	0.873	0.933	0.971	0.987	1	0.851	0.91	0.943	0.971	0.993	0.936	0.045
mean *p*	0.007	0.037	0.053	0.073	0.095	0.012	0.036	0.054	0.073	0.103	0.056	0.028
a_{effort}	0.039	0.193	0.283	0.416	0.588	0.072	0.209	0.321	0.44	0.66	0.338	0.185
N_{super}	148.6	196.6	235	302.1	522.8	183.7	230.7	281.1	365.1	672	332.5	176.7
λ	0.943	0.984	1.018	1.043	1.1	0.948	0.99	1.018	1.049	1.105	1.021	0.049
N_1	55.9	105.5	143.5	193.2	339.3	85.1	128.8	170.2	231.8	437.9	202.5	128.2
N_2	68.4	111.8	146.5	192.8	337.5	96.3	134.4	172.5	231.8	437.9	206	125
N_3	79.9	117.7	149.7	193.4	341.7	104.2	140	175.8	234.1	442.4	209.9	123.6
N_4	84.9	122.7	151.8	195.4	343.7	109.8	144.5	179.2	236.3	444.6	213.9	123.5
N_5	89.9	125.1	157.5	198.6	351.5	113.1	147.8	182.6	239.7	449.1	217.9	124.6
N_6	88.9	127.4	160	201.4	355.3	115.4	150.1	185.9	244.2	454.7	221.6	126.7
N_7	89.8	128.2	166.7	205.1	362	116.5	152.3	189.3	249.8	462.6	225.5	129.4
N_8	87.6	129.2	170.8	212.3	366.8	115.4	154.3	192.6	256.5	469.3	228.9	132.9
N_9	83.9	129.1	174.3	216.5	373.2	113.1	154.6	194.9	260.1	479.4	232.1	137
$B_{2:t}$	0	4.8	9.2	16.7	30.4	2.2	7.8	13.4	21.3	38.1	16.1	13
σ_{tRE}^2	0.00	0.00	0.04	0.13	0.52	0.00	0.05	0.13	0.29	0.78	0.23	0.36
σi_{RE}^2	0.00	0.00	0.04	0.24	1.28	0.00	0.04	0.19	0.59	1.70	0.43	0.60

Power analysis

Results for the power analysis are presented in Table 3 and Figure 6. Scenarios 1 and 2, which are representative of the current sampling program to date, show that we currently have very low power to detect population declines for the population of *Ziphius* using SOAR, at least using a frequentist implementation of a Pradel-model approach and a conventional null-rejection threshold of $\alpha = 0.05$. Scenarios 3-6 show that doubling the sample size by continuing the time series increases power faster than does doubling within-year effort, although the latter would support detection of a population trend sooner than simply extending the time series would.

Scenarios 7 to 12 showed that individual random effects make little difference to power or confidence interval coverage for key demographic rate parameters (but these would be important for unbiased population size estimation). Scenarios 13 to 18 suggest apparent higher power when less of the temporal variation is explained, but this is actually due to lower accuracy and confidence interval coverage for λ (Scenarios 13 to 18 also resulted in higher percentages of simulations for which the lower confidence limit was greater than one, erroneously concluding an increasing trend when there was a decreasing trend). This underscores the importance of accounting for temporal random effects to the best extent possible in assessing population trends.

Finally, Scenarios 19 to 21 suggest that for this study, sampling design in terms of spacing over the year or restriction to one season makes little difference in power and confidence interval coverage, so long as sampling is consistent through time. The potential for confounding population trends with secular trends in *p* indicates that maintaining a constant seasonal mix of effort or ensuring that opportunistic sampling does not vary directionally over time with respect to season is important, at least so long as we continue to have insufficient data to support detecting seasonal differences in *p* in the the mark-recapture models. This advice on sampling design should not be applied to different studies and particularly not to shorter time series.

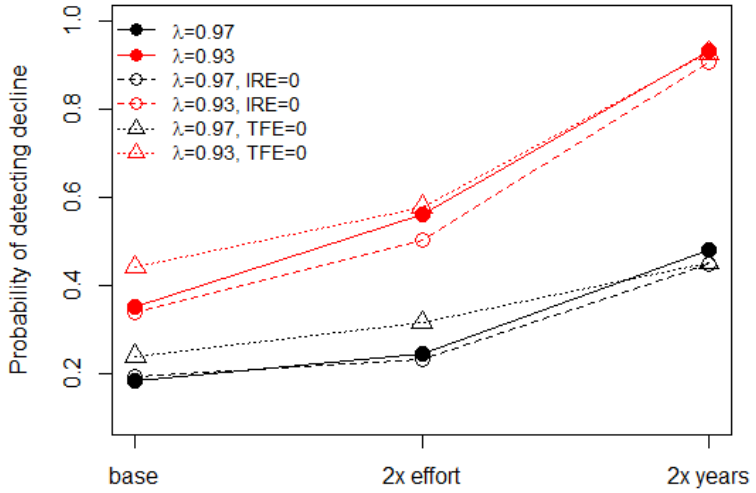

Figure 6. Probability of detecting a declining trend in Scenarios 1 to 18. Probability of detecting decline was measuerd as the percentage of simulations in which, given model convergence, the upper confidence limit (97.5th percentile) for λ was less than 1. Black points and lines represent scenarios with λ of 0.97. Red points and lines represent scenarios with λ of 0.93. Solid points and lines represent Scenarios 1 to 6 with $\sigma_{tFE}^2 = \sigma_{tRE}^2 = 0.125$ and $\sigma_{iRE}^2=0.25$. Open circles and dashed lines represent Scenarios 7 to 12 with $\sigma_{iRE}^2=0$. Open triangles and dotted lines represent Scenarios 13 to 18 with $\sigma_{tFE}^2=$ and $\sigma_{tRE}^2 = 0.25$. See "Work Completed" for full explanation of different scenarios.

Table 3. Results from frequentist power analysis based on fitting simulated capture histories with Pradel-lambda model in MARK via RMark. Scenarios are as described in "Work Completed". Performance measures include percentage simulations where model converged to produce a maximum likelihood estimate, and of those, percentage simulations where a decline was detected (i.e., upper 95% confidence limit for $\lambda < 1$), percentage simulations where true φ was within the estimated confidence interval (CI), and percentage simulations where true λ was within the estimated confidence interval.

Scenario	% Converged	% Detected decline	φ CI coverage	λ CI coverage
1: λ=0.97, 9 y, $\sigma_{tFE}^2 = \sigma_{tRE}^2 = 0.125$	77.0	18.6	94.3	85.4
2: λ=0.93, 9 y, $\sigma_{tFE}^2 = \sigma_{tRE}^2 = 0.125$	83.0	35.2	96.4	87.6
3: λ=0.97, 18 y, $\sigma_{tFE}^2 = \sigma_{tRE}^2 = 0.125$	99.9	48.2	92.8	90.4
4: λ=0.93, 18 y, $\sigma_{tFE}^2 = \sigma_{tRE}^2 = 0.125$	99.7	93.5	93.9	92.7
5: λ=0.97, 9 y, effort doubled, $\sigma_{tFE}^2 = \sigma_{tRE}^2 = 0.125$	95.3	24.6	93.2	84.7
6: λ=0.93, 9 y, effort doubled, $\sigma_{tFE}^2 = \sigma_{tRE}^2 = 0.125$	98.0	56.3	95.6	86.2
7: see Scenario 1, but σ_{iRE}^2=0	71.0	19.4	94.9	85.6
8: see Scenario 2, but σ_{iRE}^2=0	78.5	34.1	95.9	87.3
9: see Scenario 3, but σ_{iRE}^2=0	98.8	44.8	95.6	90.8
10: see Scenario 4, but σ_{iRE}^2=0	99.6	90.9	94.8	92.6
11: see Scenario 5, but σ_{iRE}^2=0	88.3	23.4	96.6	83.6
12: see Scenario 6, but σ_{iRE}^2=0	93.9	50.5	97.3	84.2
13: see Scenario 1, but σ_{tFE}^2=0, $\sigma_{tRE}^2 = 0.25$	78.4	24.1	95.7	78.8
14: see Scenario 2, but σ_{tFE}^2=0, $\sigma_{tRE}^2 = 0.25$	84.3	44.2	96.1	83.5
15: see Scenario 3, but σ_{tFE}^2=0, $\sigma_{tRE}^2 = 0.25$	100	45.2	94.9	86.0
16: see Scenario 4, but σ_{tFE}^2=0, $\sigma_{tRE}^2 = 0.25$	100	92.9	94.4	88.8
17: see Scenario 5, but σ_{tFE}^2=0, $\sigma_{tRE}^2 = 0.25$	94.7	31.7	90.9	73.3
18: see Scenario 6, but σ_{tFE}^2=0, $\sigma_{tRE}^2 = 0.25$	97.3	58.0	92.7	77.4
19: λ=0.97, 9 y at every 4 mo	76.6	22.8	93.7	82.4
20: λ=0.97, 9 y at 3 consec. mo	79.5	25.9	94.1	78.5
21: λ=0.97, 9 y at every 4 mo, increased σ_{tRE}^2, lower p	75.9	24.2	94.6	75.2

IMPACT/APPLICATIONS

Our analyses confirm the usefulness and importance of the ongoing photo-ID and passive acoustic monitoring programs for beaked whales on SOAR. The two programs collect different but complementary data. The hydrophone array provides information about daily densities of animals using SOAR, from which seasonal and annual trends in use-patterns can be inferred. The photo-ID work allows estimation of the absolute abundance of animals using SOAR on an ongoing basis (the latter is an order of magnitude larger than the former), trends in abundance, and estimates of demographic parameters such as apparent annual survival. The acoustic program can be used to detect short-term (e.g., behavioral) responses to disturbance and potentially long-term changes in use patterns, while the mark-recapture program provides information about longer-term demographic changes. Moreover, the support provided by the NUWC's M3R program to the mark-recapture program via detection and location of beaked whales with the hydrophone array also directly adds power to the mark-recapture program by increasing the efficiency of mark-recapture effort in face of low capture rates for beaked whales, and by supporting our ability to infer that SOAR is sampled fairly evenly by the mark-recapture program. The combination of the two sampling programs is clearly greater than the sum of its parts in terms of information gained.

Our current low power to detect a decline in population size from the photo-ID data can be augmented most efficiently through additional years of data. The mark-recapture time series is at a point where

each additional year of data substantially improves precision and potentially also accuracy of population and demographic rate estimates. Additional effort within each year sampled can increase power over a shorter time frame, although dollar for dollar, increasing the number of years of data-gathering will provide more information in the long term than increasing data-collection effort within years. As such, it makes sense to focus on lengthening the time series.

In a recent meeting of the Navy-funded "Population Consequences of Disturbance" (PCoD) project (held in Santa Cruz, CA, February-March 2017), the working group (in which Moore participated) extensively discussed and attempted to catalogue measurable and responsive metrics of demographic impacts caused by human activities. The working group generally agreed that, for long-lived species such as large odontocetes, demographic characteristics such as reproductive success, calf and juvenile survival, and resulting changes in age structure will be more affected in the short-term by disturbances and environmental changes than adult survival rates or population trend. Acoustic monitoring cannot provide data about these finer-scale demographic characteristics, and they can be challenging to derive from the current photo-ID data alone for some classes of individuals (most calves are insufficiently marked for recapture, and adult females can be difficult to distinguish from sub-adults in this population). While the adaptation of systematic methods to improve differentiation of age and sex classes from photos is underway (e.g., Coomber et al. 2016), increased biopsy sampling effort began in 2016 and is providing additional sex and maturity data, including pregnancy rates, for whales in this study. Continuing biopsy sampling effort throughout the extended photo-ID program may be the most effective approach for detecting changes in population health.

RELATED PROJECTS

Moretti, David

- Advanced methods for passive acoustic detection, classification, and localization of marine mammals)

Moretti, David and Falcone, Erin

- A population consequence of acoustic disturbance model for Cuvier's beaked whale (*Ziphius cavirostris)* in southern California: Photo-ID and tag data components

REFERENCES

Baumann-Pickering, S., J.A. Hildebrand, T. Yack, and J.E. Moore. 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.

Bonner, S.J. and J. Holmberg. 2013. Mark-recapture with multiple, non-invasive marks. *Biometrics* 69:766-775.

Cooch, E.G. and G.C. White. 2016. Appendix F: Parameter identifiability by data cloning. In: Cooch EG, White GC (eds) *Program Mark: A Gentle Introduction*, pp F1-F21.

Coomber F., Moulins A., Tepsich P., and M. Rosso. 2016. Sexing free-ranging adult Cuvier's beaked whales (*Ziphius cavirostris*) using natural marking thresholds and pigmentation patterns. *Journal of Mammalogy*. 97(3):879–90.

Cox, T. M., T. J. Ragen, A. J. Read, E. Vos, R. W. Baird, K. Balcomb, J. Barlow, J. Caldwell, T. Cranford, L. Crum, A. D'Amico, G. D'Spain, A. Fernandez, J. Finneran, R. Gentry, W. Gerth, F. Gulland, J. Hildebrand, D. Houser, T. Hullar, P. D. Jepson, D. Ketten, C. D. MacLeod, P. Miller, S. Moore, D. C. Mountain, D. Palka, P. Ponganis, S. Rommel, T. Rowles, B. Taylor, P. Tyack, D. Wartzok, R. Gisiner, J. Mead, and L. Benner. 2006. Understanding the impacts of anthropogenic sound on beaked whales. *Journal of Cetacean Research and Management* 7(3):177-187.

Falcone, E. A., G. S. Schorr, A. B. Douglas, J. Calambokidis, E. Henderson, M. F. McKenna, ... & D. Moretti. 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(12), 2631-2640

Falcone, E. A., G. S. Schorr. 2011. Distribution and demographics of marine mammals in SOCAL through photo-identification, genetics, and satellite telemetry: a summary of surveys conducted 15 June 2010 – 24 June 2011. Report prepared for the Chief of Naval Operations, Washington DC. Published by the Naval Postgraduate School, Monterey, California 92943 under NPS Grant N00244-10-1-0050. Available at: http://calhoun.nps.edu/handle/10945/722 (accessed 2014 Sept 26).

Falcone, E. A. and G. S. Schorr. 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, California 92943 under NPS Grant N00244-10-1-0050. Available at: http://calhoun.nps.edu/handle/10945/44226 (accessed 2016 Jan 22).

Kellner, K. 2016. jagsUI: A Wrapper Around 'rjags' to Streamline 'JAGS' Analyses. R package version 1.4.2. http://CRAN.R-project.org/package=jagsUI

Kéry, M., and M. Schaub. 2011. *Bayesian population analysis using WinBUGS*. Academic Press, Burlington.

King, R., B.T. McClintock, D. Kidney, and D. Borchers. 2015. Capture-recapture abundance estimation using a semi-complete data likelihood approach. *The Annals of Applied Statistics* 10(1). arXiv preprint arXiv.1508.06313.

Lele, S. R., B. Dennis, and F. Lutscher. 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., K. Nadeem, and B. Schmuland. 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., J. Yoshizaki, L.L. Bailey, and K.H. Pollock. 2010. Uncovering a latent multinomial: Analysis of mark-recapture data with misidentification. *Biometrics* 66:178-185.

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., P.B. Conn, R.S. Alonso, and K.R. Crooks. 2013. Integrated modeling of bilateral photo-identification dta in mark-recapture analyses. *Ecology* 94:1464-1471.

NRC 2005. Marine Mammal Populations and Ocean Noise. Determining when noise causes biologically significant effects. *The National Academy Press*, Washington DC. 126pp.

Laake, J.L. 2013. RMark: An R Interface for Analysis of Capture-Recapture Data with MARK. AFSC Processed Rep 2013-01, 25p. Alaska Fish. Sci. Cent., NOAA, Natl. Mar. Fish. Serv., 7600 Sand Point Way NE, Seattle WA 98115.

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

R Core Team. 2015. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.Royle JA, Dorazio RM. 2008. *Hierarchical Modeling and Inference in Ecology*. Academic Press, London.

White, G.C. and K.P. Burnham. 1999. Program MARK: Survival estimation from populations of marked animals. *Bird Study* 46 Supplement, 120-138.

White, G.C. and E.G. Cooch. 2017. Population abundance estimation with heterogeneous encounter probabilities using numerical integration. *The Journal of Wildlife Management* 81:322-336.

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.