

# Photographic mark-recapture analysis of clustered mammal-eating killer whales around the Aleutian Islands and Gulf of Alaska

Durban, J.<sup>1</sup>, Ellifrit, D.<sup>2,3</sup>, Dahlheim, M.<sup>1</sup>, Waite, J.<sup>1</sup>, Matkin, C.<sup>2</sup>, Barrett-Lennard, L.<sup>2,4</sup>, Ellis, G.<sup>2,5</sup>, Pitman, R.<sup>6</sup>, LeDuc, R.<sup>6</sup>, and Wade, P.<sup>1</sup>

<sup>1</sup> National Marine Mammal Laboratory, Alaska Fisheries Science Center, National Marine Fisheries Service, NOAA, 7600 Sand Point Way NE, Seattle, WA 98115, USA

<sup>2</sup> North Gulf Oceanic Society, 3430 Main Street, Homer, Alaska 99603, USA

<sup>3</sup> Center for Whale Research, 355 Smuggler's Cove Road, Friday Harbor, WA 98250, USA

<sup>4</sup> Vancouver Aquarium Marine Science Center, PO Box 3232, Vancouver, B.C., V6B 3X8, Canada.

<sup>5</sup> Pacific Biological Station, Fisheries and Oceans Canada, 3190 Hammond Bay Road, Nanaimo, British Columbia, V9R 5K6, Canada.

<sup>6</sup> Southwest Fisheries Science Center, 8604 La Jolla Shores Drive, La Jolla, California 92037, USA

Corresponding Author: [John.Durban@noaa.gov](mailto:John.Durban@noaa.gov); Tel. 206-661-5443

## Abstract

We used photographic mark-recapture methods to estimate the number of mammal-eating “transient” killer whales using the coastal waters from the central Gulf of Alaska to the central Aleutian Islands, around breeding rookeries of endangered Steller sea lions. We identified 154 individuals from 6489 photographs collected between July 2001 and August 2003. A Bayesian mixture model estimated seven distinct clusters (95% Probability Interval = 7-10) of individuals that were differentially covered by 14 boat-based surveys exhibiting varying degrees of association in space and time. Markov Chain Monte Carlo methods were used to sample capture probabilities across the distribution of clusters and estimate a total of 345 whales (95% Probability Interval = 255 – 487). Estimates of covariance between surveys, in terms of their coverage of these clusters, suggested population structure and seasonal movements from these near-shore waters, representing spatial and temporal variation in the predation pressure on coastal marine mammals.

**Keywords.** *Mark-recapture; Killer whale; Bayesian; Mixture models; Steller sea lions; Aleutian Islands; Gulf of Alaska*

## Introduction

There is growing interest in the ecosystem role of apex predators, particularly in their ability to effect top-down forcing on ecosystem dynamics through predation (Pace et al. 1999). Killer whales (*Orcinus orca*) are the top marine predator, with a world-wide distribution (Dahlheim and Heyning 1999). In the far North Pacific waters of the western Gulf of Alaska, Aleutian Islands and Bering Sea, predation by killer whales has recently been suggested as a possible cause for the declines in abundance of several marine mammal species, notably endangered Steller sea lions (*Eumetopias jubatus*) and sea otters (*Enhydra lutris*) (Estes et al. 1998; Springer et al. 2003; Williams et al. 2004). Although there is considerable disagreement over the evidence supporting this hypothesis (Demaster et al. 2006; Mizroch and Rice 2006; Trites et al. 2007; Wade et al. 2007; Springer et al. 2008; Estes et al. 2009), this debate has highlighted the need for additional empirical data on the role of killer whales as predators within these marine ecosystems.

Killer whales in the North Pacific have been categorized into sympatric lineages that display differences in genetic composition (Stevens et al. 1989; Hoelzel and Dover 1990; Hoelzel et al. 1998; Barrett-Lennard 2000; Hoelzel et al. 2002), morphology (Baird and Stacey 1988; Ford et al. 2000), vocalizations (Ford 1989; Barrett-Lennard et al. 1996; Deecke et al. 2005) and social structure (Bigg et al. 1990; Baird and Whitehead 2000; Parsons et al. 2009). These lineages function as separate ecotypes and exhibit distinct prey specializations, with “residents” and “offshores” feeding on fish, and “transients” on marine mammals (Ford et al. 1998; Saulitis et al. 2000; Herman et al. 2005; Krahn et al. 2007) Owing to these prey preferences, the different killer whale ecotypes will have differing roles in the dynamics of marine ecosystems. However, useful inference about ecosystem dynamics has been constrained by the absence of precise abundance estimates for killer whales (Guenette et al. 2006).

Killer whales can be individually recognized from natural markings, and in several areas of the northeast Pacific accessible populations of fish-eating residents have been enumerated in photo-identification censuses (Dahlheim et al. 1997; Matkin et al. 1999; Ford et al. 2000). Photo-identification catalogues have also been compiled for the mammal-eating transients (Black et al. 1997; Dahlheim et al. 1997; Ford and Ellis 1999; Matkin et al. 1999) but the difficulties of encountering transients on a regular basis have prevented precise counts of abundance. Furthermore, although killer whales have been relatively well documented for the coastal waters of the northeast Pacific, relatively few data exist for the more remote waters of the far North Pacific, requiring a sampling approach to abundance estimation (Zerbini et al. 2007). Here we estimated the abundance of mammal-eating killer whales using the coastal waters from the central Gulf of Alaska to the central Aleutian Islands around the breeding rookeries of the endangered western stock of Steller sea lions (Loughlin et al. 1992). Our estimate was based on photo-identification samples that utilized natural markings to identify individual killer whales, and we applied mark-recapture models to these data to estimate the number of individuals that remained undetected.

Mark-recapture methods are well developed for estimating the size of wildlife populations (Chao 2001) and photographic “mark-recapture” methods have been used to estimate the abundance of a number of populations of whales and dolphins (Wilson et al. 1999; Stevick et al. 2003; Calambokidis and Barlow 2004; Durban et al. 2005). In the traditional ecological framework, the researcher controls the capture mechanism and thus the individual probability of being caught (Otis et al. 1978; Seber 1982). However, when working with cetaceans in remote marine environments, which are costly and difficult to survey, the design is generally more observational. Surveys may be setup for various purposes, and the researcher cannot easily control the probability of inclusion in a photo-

identification sample. Additionally, cetacean populations tend to be structured by complex social organization, with heterogeneous ranging patterns between individuals (Lusseau et al. 2006). Therefore, depending on their spatial and temporal arrangement, the survey samples tend to cover different subsets of the population, with some overlap (Durban et al. 2005).

Killer whales present a special case of heterogeneity because they typically travel in matrilineal groups that can remain stable in composition for long periods of time, on the scale of years for transients (Baird and Whitehead 2000) to generations for residents (Parsons et al. 2009). Therefore, heterogeneity is likely to be represented by groups of whales, each with different capture probabilities by different surveys. This clearly complicates the relationship between recapture rate and abundance, and has the potential to produce biased estimates of abundance using conventional mark-recapture models (Carothers 1973a; Carothers 1973b; Otis et al. 1978). For our study, we tailored mark-recapture models to fit the key interactions between clustered capture probabilities and non-random surveys.

## Methods

### Survey Data

Photo-identification data were collected within an area of approximately 220,000 km<sup>2</sup> of the near-shore waters of the Aleutian Islands and western Gulf of Alaska (Fig. 1). This area was designed to cover much of the known haul-out range of the endangered western stock of Steller sea lions in U.S. waters (Loughlin et al. 1992) from the Kenai Peninsula in the east (~60°N, 150°W) to Amchitka Pass in the central Aleutians (~52°N, 178°W). In summer, Steller sea lions are thought to primarily forage in relatively close proximity to their rookeries and haulouts (Merrick and Loughlin 1997) and therefore our surveys covered the area within 60 km of land. Photo-identification surveys were conducted over a 26 month period between July 2001 and August 2003, with survey effort confined to the summer months between May and August in each year. Our abundance estimates correspond to the number of whales that used this study area during these times, and we make no explicit assumption as to how this estimate relates to the total population to which these animals belong. Also, it is not assumed that all of these individuals were present in the area for the entire survey period.

Photo-identification data were obtained from 14 different surveys (Table 1), differing in their spatial and temporal coverage within the survey area. The widest geographical coverage was obtained during three line-transect surveys (01LT, 02LT, 03LT), one each summer (Zerbini et al. 2007), which defined the extent of the overall survey area. These were dedicated surveys to assess killer whale distribution and abundance, and employed large ships to sample the entire study area following track-lines that were randomly placed in a saw-tooth pattern. The ship diverted from the track-line to photograph any killer whale groups that were sighted. To ensure equal area-coverage for the line-transect methodology, the survey area was stratified into 16 strata (Fig. 1), which are useful here for describing the spatial coverage of the other surveys (Table 1).

Dedicated killer whale surveys were also conducted as part of focused research efforts in two near-shore areas: one in the eastern Aleutian Islands (strata 10, 11 and 12), and one in the western Gulf of Alaska between the Kenai Fjords and Kodiak (strata 1, 2 and 3). Surveys in each year (01EA, 02EA, 03EA, 03FP in the eastern Aleutians; 01KK, 02KK, 03KK in the Kenai/Kodiak region) were conducted using chartered fishing vessels (Matkin et al. 2007). Daily or multi-day research trips were repeatedly conducted over several weeks in each summer, and survey effort was targeted in areas that killer whales were known or expected to

use. Killer whale photographs were also collected on an opportunistic basis from other oceanographic and biological research platforms that were operating in the study area in the summer of 2002: 02AH (Sinclair et al. 2005), 02MF (Waite et al. 2002); 02MA (LeDuc 2004); and 02TX (Sease and Gudmundson 2002) (Table1).

On all surveys, killer whale identification photographs were taken using 35mm SLR cameras equipped with Fuji Neopan 1600 black and white film, or with digital SLR cameras shooting high resolution images. To standardize data collection, only the left side of each whale was photographed, specifically the dorsal fin and adjacent saddle patch that are individually distinctive (Fig. 2), and all the individuals in each group were photographed when possible.

### Mark-recapture samples

The focus of this study was to estimate the abundance of transient killer whales. However, most of the killer whale groups encountered within this study area were of the resident lineage (Zerbini et al. 2007). We therefore needed to identify groups of mammal-eating transients before using data from only these groups in the subsequent analyses. Identification to lineage was based on a combination of molecular genetic analysis of biopsy samples and photographic assessment of morphology (Matkin et al. 2007; Zerbini et al. 2007). Chemical analyses of fatty acids, stable isotopes and organochlorines from the same biopsy samples have demonstrated that the animals genetically identified as transients in this area had chemical signals that were consistent with a diet of marine mammals (Herman et al. 2005; Krahn et al. 2007). This dietary specialization is also supported by direct observations of transients preying solely on marine mammals in this area (Matkin et al. 2007).

To minimize erroneous documentation of markings (i.e. individual identity), all photographic identifications were subject to a quality grading process. Because identifying features on both the dorsal fin and saddle patch were used, this grading scheme reflected the quality of both of these body parts in the image. Specifically, a grade of 4 indicated that the both the fin and saddle were displayed in usable quality; 3 indicated a usable fin only (e.g., animal not surfacing high enough to display most of the saddle); 2 indicated a usable saddle only (e.g., the dorsal fin was not framed in the image); and 1 indicated that neither feature was usable. The utility of the images was assessed based on a combination of image size, focus, angle and clarity due to image exposure and was judged by whether it could be used to document the smallest of distinguishing marks on a hypothetical individual.

The mark-recapture procedure required natural markings that were sufficiently distinct to allow for reliable repeat identifications over time. Therefore, we scored the markings present on each individual to determine which were distinctively marked. Specifically, each photograph was scored for the presence or absence of five key features: distinctive dorsal fin shape, notches in the dorsal fin, distinctive saddle pigmentation pattern, scratches on the saddle and oval scars on the saddle (Fig. 2). These marks were classified as either primary or secondary, relating to their utility for defining individual distinctiveness. Notches in the dorsal fin were considered to be a primary mark, as they have been shown to persist with little changes for many years in free-ranging killer whales (Ford et al. 2000). The size, shape and position of notches in the dorsal fin make them a distinguishing feature that can be reliably documented even in a wide variety of photographic lighting conditions. The other mark-types were conservatively considered to be secondary marks because of uncertainty over their longevity and because their reliable documentation is conditional on the light quality in the photographic image. Specifically, many of the photographed whales possessed oval scars on the saddle patches that are thought to be caused by cookie-cutter sharks, *Isistius* sp. (Durban and Pitman, unpublished data). These are deep scars that will

most likely be visible for at least several years. However, even when controlling for light levels in the quality grading procedure, the detail of these scars can be relatively hard to see in low and flat light. Because of the subjectivity involved, guidance about which whales possessed distinctive fin and saddle shapes was based on soliciting independent expert opinion from a number of the authors who are experienced in viewing killer whales (GE, DE, MD, JW, JD).

Individuals were defined to be distinctively marked if they possessed either a minimum of one primary mark or two secondary marks, and only distinctively marked individuals were used in the mark-recapture analysis. Each individual was assigned a unique identification number, and individual identity was assigned by comparing candidate photographs to an updating catalogue of unique individuals. However, because repeated photographs of the same individual varied in photographic quality, the capture histories for each whale were constructed based on a combined query of both the quality grades for each photograph and the individual distinctiveness scores, to minimize identification errors. Specifically, an identification was only considered usable if the one primary or two secondary distinctive features were documented on a part of whale (dorsal fin, saddle or both combined) that was graded to be usable quality.

### Mark-recapture models

We constructed a matrix of identification histories  $X_{S \times n}$  with elements  $x_{ij}$  taking the values 1 or 0 to indicate whether or not each individual  $j = 1, \dots, n$  was identified during each of the  $i = 1, \dots, S = 14$  surveys. These binary data were assumed to constitute the outcome of independent Bernoulli trials, conditional on capture (or in this case “identification”) probabilities  $p_{ij}$ . The aim of mark-recapture abundance estimation was to predict the number of individuals that remained undetected, using information on the  $n$  observed individuals. This was achieved by augmenting the observed data matrix to include  $n_0 = N - n$  rows of zeros corresponding to the unobserved individuals from a total population of size  $N$ , with the dimension of  $n_0$  being predicted directly from the model for the observed data by assuming that the same probability for unobserved individuals  $p_{ij}, j = n + 1 \dots n + n_0$ , as for observed individuals  $p_{ij}, j = 1 \dots n$  (Durban and Elston 2005). Emphasis was therefore placed on models to describe the observation process generating the identification probabilities.

A consequence of our non-uniform and opportunistic survey effort was that the probability of identification likely varied between individuals and across sampling occasions. We assumed the population was demographically closed over this two-year period, and modeled variability in identification probability through just two types of main effect: survey effects and individual effects. These effects were incorporated into a linear logistic model for the identification probability (Pledger 2000) where the most general model takes the following form:

$$\log \{p_{ij}/(1 - p_{ij})\} = \mu + \alpha_i + \theta_j + \lambda_{ij}$$

where  $\mu$  is an overall constant representing the average identification probability,  $\alpha_i$  is the main effect depending on the survey  $i$ ,  $\theta_j$  is an individual effect and  $\lambda_{ij}$  represents the interaction of individuals across surveys. The general model, labeled  $M_{\alpha\theta\lambda}$ , could be

simplified to form a variety of alternative model forms. If we removed the interaction term  $\lambda$ , this became an additive model where the individual effects were constrained to vary in a parallel fashion across surveys ( $M_{\alpha\theta}$ ). We could further simplify the model by removing either or both of the main effects terms to result in a model depicting only survey ( $M_{\alpha}$ ), or individual ( $M_{\theta}$ ) variability in identification probability, or no systematic variability around an average identification probability ( $M_0$ ). The full model, and its various sub-models are listed in Table 2.

## Hierarchical Bayesian formulation

To ensure parameter identifiability, we adopted a random effects approach for modeling these parameters by assuming an underlying common distribution to describe variability in effects across surveys, individuals and their interaction (Coull and Agresti 1999; Dorazio and Royle 2003). This random effects formulation was specified within a Bayesian statistical framework, where the underlying common densities were conveniently specified by prior distributions (Fienberg et al. 1999; Durban and Elston 2005). We modeled the individual effects as clustered effects by adopting a mixture model for  $\theta$ , where the underlying density of the effects was approximated by a mixture of distributions (West 1992). Specifically,  $\theta$  was assumed from a mixture of  $C$  Normal component distributions:

$$\theta_j = \sum_{k=1}^C w_k N(0, \sigma^\theta).$$

The  $w_k$  were mixture weights, which summed to 1, indicating the probability of belonging to mixture,  $k = 1, \dots, C$ . Instead of a separate effect for each individual, the model therefore specified a separate effect for clusters (mixtures) of individuals. The mixtures themselves were defined by hyper-parameters specifying the mean and standard deviation of the Normal components. The use of the overall mean  $\mu$  in the additive logistic model conveniently resulted in centering of the effects  $\theta$  around zero, and the component means could therefore be set at zero. We assigned a common standard deviation hyper-parameter to describe the variability between the components' effects  $\sigma^\theta$ , but to allow non-zero effects to emerge  $\sigma^\theta$  was assigned a Uniform prior distribution between 0 and 10. Examination of later results shows that this prior standard deviation was much greater than the corresponding posterior standard deviation.

We also modeled the individual by survey interactions as clustered effects. However, because surveys were typically not randomly or uniformly placed in space or time, we also attempted to describe possible dependencies between the surveys in terms of similarities in the specific clusters that were covered by each survey. To describe the association between the 14 surveys, the prior distribution for  $\lambda$  was taken to be a multivariate Normal distribution, stratified into  $q = 14$  dimensions (Fienberg et al. 1999):

$$\lambda_{ij} = \sum_{k=1}^C w_k N_{14}(0, \Sigma)$$

This distribution was defined by the covariance matrix  $\Sigma$  of the order  $q \times q$ , where the principal (left to right) diagonal element of this matrix was the estimate of the variance of the cluster effects for each of the 14 surveys, and the off-diagonal values represented covariances between pairs of surveys. A common covariance matrix was thus specified across clusters in order to learn about the covariances between surveys in terms of how they captured the effects of each cluster. For example, a cluster with high catchability in survey 1 would have high catchability in survey 3 if a positive covariance existed. We adopted the Wishart distribution as the prior for the inverse covariance matrix  $\Sigma^{-1}$  (Fienberg et al. 1999), which is specified in terms of a scale matrix  $B$  and a degrees of freedom parameter  $\nu$ . We set diagonal values of  $B = 1$  for the prior variance of the cluster catchability for each survey  $q$ , and the off-diagonals were assigned  $B = 0$  for a prior assumption of no covariance between pairs of surveys. Larger values of  $\nu$  represent stronger belief, and we therefore adopted a value of  $\nu = q = 14$  to represent a vague prior and allow non-negative covariance values to emerge.

The allocation of individuals to clusters was not known and we instead took the approach of estimating the number of non-empty mixtures from a ceiling value by adopting a Dirichlet process prior (Escobar and West 1995; Dey et al. 1999). For a ceiling of  $C$  potential mixtures, a set of  $C$  potential values of  $\theta_k$  and  $\lambda_{ik}$  for  $k = 1, \dots, C$  were drawn from the baseline priors, with the size of the set corresponding to the anticipated maximum number of possible clusters. The most appropriate values  $\theta_k$  and  $\lambda_{ik}$  for each individual  $j$  was then selected using a Dirichlet vector of length  $C$ , with uniform elements for the mixture weights  $w_k$ . Because some of the  $C$  clusters may not have been appropriate for any of the observations (i.e., there may be empty clusters), the mixture weights  $w_k$  were chosen by the “stick-breaking” construction (Congdon 2003):

$$w_1 = r_1; w_k = (1 - r_1)(1 - r_2) \dots (1 - r_{k-1})r_k \quad k = 2, \dots, C$$

where  $r_1, r_2, \dots, r_{C-1}$  are random variables with prior mass equally spaced between zero and one. Specifically we assigned  $r_1, \dots, r_k$  with  $B(1,1)$  priors, where  $B(a,b)$  indicates a Beta distribution with mean  $m = a / (a + b)$  and variance  $v = m(1 - m) / (a + b + 1)$ . We set  $r_C = 1$  to ensure that  $\sum_{k=1}^C w_k = 1$ . This is known as a stick-breaking prior since at every stage what is left of a stick of unit length is broken, and the length of the broken portion is assigned to the current value  $w_k$  (Dey et al. 1999).

In addition to the clustered individual effects, we also set prior distributions on the other terms in the model. We adopted a random effects formulation for the survey effects  $\alpha$ , with a continuous Normal distribution describing the variability in the degree to which each survey penetrated into the target population. As with the individual main effects, the prior distribution for the survey effects was also centered on zero with standard deviation  $\sigma^\alpha$  uniformly distributed between 0 and 10. Additionally, the overall mean level of capture probability,  $\mu$ , was assigned a vague Normal prior centered at zero with a large variance ( $=100$ ).

In a Bayesian analysis the prior distributions are updated, conditional on the observed data, to produce a “posterior” probability distribution (Gelman et al. 1995). We used computer-intensive Markov Chain Monte Carlo (MCMC) sampling to perform this updating and estimate a sample from the posterior distribution for each parameter of interest (Durban and Elston 2005). To facilitate MCMC computation in the mixture models, we introduced

indicator variables  $z_j$ , one for each individual  $j = 1, \dots, n$ , each indicating which component in the mixture had generated the effect for each individual (Neal 2000). These cluster indicators were stochastic and modeled as latent variables. Conditional on the mixture weights  $w$ , each component of the indicator vector  $z$  was independently drawn from the multinomial distribution with  $\text{pr}(z_j = k) = w_k$ . The indicators therefore took the values of  $k = 1 \dots C$ , such that  $z_j = k$  indicated that the  $j$ th individual had been drawn from the  $k$ th component of the mixture. With repeated iterations in an MCMC chain, the average cluster allocation  $z_j$  attaching to individual  $j$  was taken over the candidate values assigned at each iteration of the chain, and variability in the sampled values represented uncertainty about mixture (or cluster) membership. The cluster indicators  $z_j$  for the unobserved individuals,  $j = n+1, \dots, n_0$ , were drawn from the same Dirichlet process as the observed individuals. MCMC sampling across the distribution of candidate clusters therefore had the effect of smoothing the capture probabilities of observed individuals, and then using this smoothed distribution to predict the number  $n_0$  of unseen individuals. Variability in  $n_0$  was specified through a prior distribution for  $N$ , and the two were linked by the logical relationship  $n_0 = N - n$ . A discrete uniform distribution was adopted as the prior for  $N$ , with support over the interval between  $n$  and  $M$ . The posterior probabilities of discrete values for  $n_0$  and thus  $N$  were simply estimated from the relative frequency of different values in the MCMC sample. This posterior distribution not only indicates the most likely values, but also allows a crucial assessment of the associated uncertainty.

## Model Selection and Evaluation

We employed the same MCMC simulation approach to generate predictive observations to compare the fit of the competing models using a posterior predictive criterion (Gelfand and Ghosh 1998). For each model, we predicted a new set of data ( $X_{new}$ ) of the same dimensions as the observed data ( $i$  in  $1, \dots, S$  and  $j$  in  $1, \dots, n$ ) by generating samples directly from the posterior distributions of the fitted model parameters. We then calculated a loss function that measured the discrepancy between the observed data,  $X$ , and the predicted data,  $X^{new}$ . As a loss function, we used the Mean Square Predicted Error (MSPE):

$$\text{MSPE} = \sum_{i=1}^S \sum_{j=1}^n [x_{ij}^{new} - x_{ij}]^2$$

As with other model selection methods, the predictive criterion achieves a compromise between the goodness-of-fit and a penalty for the number of free parameters in the model (Gelfand and Ghosh 1998). The model with the smallest criterion value was estimated to be the model that would best predict a replicate dataset of the same structure as that currently observed. However, the predictive model selection criterion did not reveal whether the selected model could have plausibly generated the observed data. We, therefore, also adopted a posterior predictive approach for goodness-of-fit checking (Gelman et al. 1996) by calculating a discrepancy for both the observed data,  $X$ , and the predicted data,  $X^{new}$ . Specifically, we adopted an individual-specific discrepancy measure  $d_j$  (Berkhof et al. 2003) to check whether the capture histories of each individual  $j$  were well-fitted by the probability model  $p$ :

$$d_j(X) = \sum_{i=1}^S |x_{ij} - p_{ij}|$$

$$d_j^{new}(X^{new}) = \sum_{i=1}^S |x_{ij}^{new} - p_{ij}|$$

We assessed the magnitude of the discrepancies by comparing the posterior distributions of each  $d_j^{new}$  with the corresponding  $d_j$  for each individual  $j$  in  $1, \dots, n$ .

For each model, we used the WinBUGS software (Lunn et al. 2000) to simulate three simultaneous MCMC sequences from the posterior distribution for the model parameters, and the predictive distribution given these posteriors. We compared three MCMC sequences in order to identify and discard “burn-in” iterations prior to convergence of the chains (Brooks and Gelman 1998), and inference was based on 10,000 iterations following the burn-in. We adopted a ceiling of  $C = 30$  possible clusters when fitting the mixture models and  $M = 500$  potentially unobserved individuals, which was notably larger than the upper bounds for the respective posterior estimates.

### Rescaling for non-distinctive individuals

Even with the use of high-quality photographs that allowed identification of individuals based on subtle markings, there remained some individuals that were not judged to be distinctive, even from a best quality ( $Q = 4$ ) photograph (Fig. 2). Reliable identification histories over time could therefore not be constructed for such individuals, so they were not included in the mark-recapture analysis. However, it was necessary to have some estimate of the proportion of these individuals in the population to re-scale the mark-recapture estimate upward to account for these non-distinctive animals.

From high quality ( $Q = 4$ ) photographs, all individuals could be distinguished from the other group members in a given encounter, based on at least one secondary mark (Wilson et al. 1999). Therefore, even though some individuals could not be matched over time, we could produce an estimate of this non-distinctive proportion of the population from the average number of individuals with and without distinctive markings documented from  $Q = 4$  photographs in each encounter. To account for the variability in this estimated proportion, we incorporated this mark-type rescaling step into the same Bayesian probability model as the mark-recapture estimation and used MCMC to sample jointly across both components. Specifically, the number of distinctly marked individuals was treated as a binomial sample from the total individuals documented from  $Q = 4$  photographs in each encounter, where the binomial probability represented the proportion of distinctive individuals ( $\pi$ ). A common proportion,  $\pi$ , was adopted to estimate the average across encounters, and we adopted a flat Beta(1,1) prior distribution with probability mass equally spaced between 0 and 1. The mark-recapture and mark-rescaling components were linked to form a single probability model, by defining the overall abundance  $P$  to equal  $N/\pi$ , and integrating the distribution to obtain discrete integers of  $P$ .

## Results

Photo-identification data were collected from 63 encounters with transient killer whales in 9 of the 16 strata within this study area (Fig. 1; Table 1). These encounters were

located across the full extent of the study area, but most were concentrated around the eastern Aleutians Islands (strata 10, 11 and 12). Assignment to the “transient” lineage was corroborated by genetic sequences from tissue samples directly collected from 52 (83%) of these encounters, with a further 5 encounters being designated through indirect association of individuals with directly sampled transient groupings. There were only six encounters from which these molecular tools could not be used, and determination was based solely on photographic assessment. Two independent analysts showed complete agreement in their classifications in these six cases, and for all photographed groups in general. Furthermore, these photographic assignments were always consistent with genetic determinations for the 57 encounters where genetic determinations were also available.

The 14 surveys varied in the quantity of data collected (Table 1). Targeted surveys (e.g., 03FP) were most effective at encountering whales, whereas the opportunistic surveys (e.g., 02MF) often only produced sparse data, and the dedicated line-transect (LT) surveys were successful in encountering whales over a wider geographical area. In total, 4932 photographs were obtained, displaying 6489 identification images (as there were often images of multiple whales identified within a single photographic frame). Just under half (2735; 42%) of the images were judged to be of sufficient quality to attempt individual identifications, and 1402 (51%) of these were high quality photographs displaying both the fin and saddle patch.

The number of distinct individuals identified from each survey also varied widely, reflecting the differences in encounter rates and photographic effort (Table 1). There were 203 individual identifications in total, the majority (129; 64%) from primary markings (notches in the dorsal fin), with 74 (36%) from a combination of two or more secondary markings. Although individuals were typically only identified from a single survey, there was some survey overlap, with 18, 11 and 3 individuals being identified from 2, 3 and 4 different surveys, respectively. No individuals were documented on more than four surveys. As a result, we identified a total of 154 different distinctly marked individuals. The rate of discovery of new individuals gradually decreased over the study period as individual matches between surveys were documented (Fig. 3). However, this rate did not level off to a stable asymptote, and there was clearly scope for undetected individuals.

There were marked differences in the fit of the five models to the observed data (Table 2). Adding clustered individual effects alone to the base model produced no improvement in model fit ( $M_0$  vs.  $M_\theta$ ), and individuals were assigned to only a single most likely cluster, reflecting little evidence for clustered effects in individual main effects under this model. Adding survey effects alone did result in a slightly better fitting model ( $M_0$  vs.  $M_\alpha$ ). This reflected the widely different coverage provided by each survey, both in terms of geographic area and the intensity or duration of survey effort. Incorporating clustered individual effects in an additive fashion to survey effects did not produce any further improvement in model fit ( $M_\alpha$  vs.  $M_{\alpha\theta}$ ), as once again the individuals were assigned to only a single cluster under this model. This demonstrated very little support for parallel variation in identification probabilities of different individuals over the 14 surveys. However, the full model ( $M_{\alpha\theta\beta}$ ) incorporating survey effects, cluster effects, cluster by survey interactions and associations between surveys clearly predicted the observed data much more accurately than any of the alternative models, with less than a third of the predictive error of any of the other candidate formulations. Therefore, we based further inference on estimates from this model.

A maximum of 10 different clusters were sampled during the MCMC run for model  $M_{\alpha\theta\beta}$ , with a modal (most probable) value of 7 and 95% probability intervals of 7-10 for the distribution of the number of non-empty clusters. Of the 154 distinctly marked individuals, 134 (87%) could be assigned to one of these seven specific clusters with high probability, with this cluster accounting for at least 95% of the discrete probability in the estimated

posterior distribution for the cluster indicator  $z$ , and 152 (99%) could be assigned to a dominant cluster with  $\text{pr}(z_j = k) > 0.5$ . This provided strong evidence for the existence of distinct groups or subpopulations, and demonstrated the utility of allowing for non-parallel variability in the catchability of these different clusters by different surveys. In general, the individual identification histories were well-fitted by the distributions of the corresponding mixture components, as the summed discrepancy value for the observed data for any individual was never larger than 3.04 (out of a possible 14 binary identification events), and averaged less than one (mean = 0.91, SD = 0.87). Furthermore, the difference between the mean discrepancy value and the mean discrepancy value predicted under the model only averaged 0.18 (SD = 0.20), implying that the data could have been plausibly generated by the model.

The estimated covariances for the cluster by survey interaction terms indicated which surveys were positively or negatively associated in relation to the clusters that they covered (Table 3). Of the 84 different pair-wise combinations of surveys, 51 (61%) had significant associations where the 95% probability intervals of the posterior distribution for the covariance  $\Sigma$  did not overlap with zero (no correlation). Of these, 27 (53%) represented negative associations and 24 positive correlations. The strongest covariance was negative, with a median estimate of -178 (95% Probability Interval = -320 to -105) for the association between the 03EA and 03FP surveys. In fact, 7 of the 13 covariances involving the 03FP survey were significant and all of these were negative. These negative associations were the result of most of the individuals identified by the 03FP survey (52 out of a total of 59) not being documented by any of the other surveys, and therefore these 52 individuals were assigned to the same cluster with at least 95% probability, which represented the largest cluster estimated by the model. The 03EA survey was also involved in the strongest positive covariance (median = 137, 95% Probability Interval = 53-235) with the 02MF survey. This is not surprising because both only had encounters in the same stratum (#11, Fig. 1). These surveys also had strong positive associations with the 02TX and 03LT surveys that also had encounters within this stratum, and these patterns defined a cluster of six individuals with high catchability across each of these surveys. A further cluster of 14 individuals was comprised of individuals that were also all identified in 03EA survey, with corresponding high catchability and positive associations to the 02EA and 02TX surveys.

Interestingly, the focused killer whale surveys in the Kodiak Island and Kenai Fjords region of the Gulf of Alaska (01KK, 02KK, 03KK) were involved in 21 (41%) of the significant associations. These were comprised of positive associations between the three KK surveys, but negative associations to any of the other surveys with encounters in the Aleutians Islands. Of the 12 individuals that were documented in the KK area, none were documented in any of the surveys elsewhere and therefore all 12 were assigned to same cluster with high catchability on the 01-03KK surveys and low catchability in others. The remaining three clusters were of similar size, ranging from 15 to 19 individuals. Two of these clusters comprised individuals that were identified by only one survey: 02LT (15 individuals) and 02AH (19 individuals), which generally had negative associations with other surveys. Finally, a cluster of 16 individuals was defined by identifies primarily in the 01LT survey, with all the individuals identified in this survey. However, there was some overlap of individuals with the 02EA, 03FP and 03LT surveys, corresponding to the positive (but weak associations) between 01LT and these surveys.

Without the cluster by survey interactions and covariance between surveys, the models with cluster effects ( $M_\theta$  and  $M_{\alpha\theta}$ ) produced only slightly higher abundance estimates than the corresponding simpler formulations ( $M_0$  and  $M_\alpha$ , respectively; Table 2). This heterogeneity effect was small because of the limited support for more than one cluster in these formulations. In contrast, a larger effect was seen through the addition of survey effects,

which resulted in decreases in abundance estimates ( $M_0$  vs.  $M_\alpha$ :  $M_\theta$  vs.  $M_{\alpha\theta}$ ) indicating the importance of accounting for variability in coverage across surveys. However, an even larger decrease in estimates was achieved when the constraint of parallel survey effects across clusters was removed through the addition of cluster by survey interactions that allowed differential coverage of specific clusters by specific surveys. Because negative associations comprised the majority, and the strongest, of the significant associations between surveys, the  $M_{\alpha\theta}$  model resulted in the lowest estimate for the abundance of distinctively marked individuals (median = 275, 95% Probability Interval = 210 – 372). Note that this estimate was greater than the  $n = 154$  individuals that were actually observed, and less than the  $M = 500$  potentially unobserved individuals that were considered, with an estimated  $n_0 = 121$  (95% Probability Interval = 56 - 218) unobserved distinctive individuals.

There were some individuals that were not judged to be distinctive, even from a high quality photograph that displayed the dorsal fin and saddle patch in good clarity (Fig. 2). These were typically younger individuals that had not yet acquired a dorsal fin notch or two types of secondary markings. We estimated that the average proportion  $\pi$  of individuals with distinctive markings per encounter was 0.80 (95% Probability Interval = 0.67 – 0.91), and therefore the abundance estimate for distinctive individuals was re-scaled to an overall abundance estimate of  $P = 345$  (95% Probability Interval = 255 – 487) (Fig. 4).

## Discussion

A recurring theme in mark-recapture studies is the need to account for heterogeneity in capture probabilities. It is well-established that behavioral differences among individuals are likely to produce variable probabilities of capture, and failure to account for this heterogeneity can lead to biased population estimates (Carothers 1973a; Carothers 1973b; Otis et al. 1978). Our analyses demonstrated that important variation in identification probabilities resulted from the use of photo-identification data from killer whales encountered during opportunistic and non-systematic surveys that induced non-equal coverage of a large survey area. Combined with differential ranging patterns of different killer whale groups, this led to the situation where clusters or “sub-populations” were differentially covered by different surveys. Rather than pooling surveys into mark and recapture samples to reduce heterogeneity (Calambokidis and Barlow 2004), we chose instead to explicitly model both heterogeneity and dependence to learn from all the information available in the overlap of individuals between surveys.

We modeled the clustered heterogeneity through the use of Bayesian mixture models, where the underlying density of individual identification probabilities was approximated by a mixture of distributions (West 1992), defining clusters of individuals with similar identification histories. Additionally, dependencies between surveys that were differentially associated in space and time were estimated using a hierarchical formulation for cluster-by-survey interaction terms (Fienberg et al. 1999), where the covariance matrix provided an estimate of the association between each pair of surveys in terms of the clusters that they covered. A key feature of our approach was that we did not specify the number of clusters, but instead estimated the number of non-empty mixtures from a ceiling value using a Dirichlet Process Prior (Escobar and West 1995). Model fitting was accomplished using MCMC sampling within a Bayesian statistical framework, allowing uncertainty in mixture allocation to be propagated into a full probability distribution for population size that effectively communicated both the extent and shape of the uncertainty (e.g., Wade 2000). The resulting abundance estimate of 345 transient killer whales (95% Probability Interval = 255 – 487) represents a key input parameter for ecosystem models (Guenette et al. 2006). For

example, energetic calculations have suggested that fewer than 40 killer whales could have caused the recent Steller sea lion decline in the Aleutian Islands, and a group of five individuals could have accounted for the decline in sea otters, if their predation was focused entirely on these species (Williams et al. 2004). It is notable that even the lower 95% probability bound of our estimate was considerably higher than this. However, our estimates of covariance between pairs of surveys suggest spatial segregation and differential site-fidelity of whales across the study area: key features of population structure need to be considered when evaluating affect the intensity and distribution of predation pressure.

We estimated significant positive associations between the three surveys in the western Gulf of Alaska (01-03KK, strata 1, Fig. 1) and negative associations between these surveys and those in the Aleutian Islands, with no overlap of individuals. Similar structure may exist between the eastern and central Aleutian Islands, as none of the individuals documented around the central Aleutians (strata 13:16, Fig. 1) have been sighted in the relatively large number of encounters in the adjacent eastern Aleutian Islands (strata 9-12). However, the absence of focused survey effort and the low sample size of encounters in the central Aleutians constrained our power to identify structural boundaries between these areas.

In contrast, we recorded the largest number of killer whale encounters in the eastern Aleutians area. Although this aggregation of sightings was partly a function of greater survey coverage, a high density of sightings in this area was also reported from randomized line-transect surveys (Zerbini et al. 2007). The strong positive associations that we estimated between surveys with encounters in the eastern Aleutians indicated the repeated re-identification of individuals, implying a relatively high level of site fidelity across different surveys and years. This apparent hot-spot in distribution is likely linked to the availability of prey. The Aleutian Passes, such as Unimak and Samalga, are highly productive regions that provide important foraging areas and movement corridors for both cetaceans and pinnipeds (Sinclair et al. 2005). Recent observations during the summer months have documented predation by killer whales on a number of pinniped and cetacean species in this area (Matkin et al. 2007), with the highest number of observations involving predation on northern fur seals (*Callorhinus ursinus*) that are increasing in number in the area due to the growth of the Bogoslof rookery (Towell et al. 2006). This area also has a high concentration of Steller sea lions, which have been increasing in number since at least 2000 (Fritz and Stinchcomb 2005).

Interestingly, the strongest negative association was estimated between the 03FP and 03EA surveys, which were geographically adjacent in the eastern Aleutians. There were also strong negative associations between the 03FP survey and all other eastern Aleutian surveys. The 03FP survey was unique in occurring earlier in the year (May 2003) and documented a large aggregation of killer whales that were feeding on young gray whales migrating northwards into the Bering Sea (Barrett-Lennard et al. In Press.). In contrast, the other eastern Aleutian surveys involved encounters from June to August after most of the gray whales had moved through the area. The majority (52 out of 59) of the individuals photographed during the 03FP survey were not identified on any other survey, suggesting that most of the animals in this late-spring aggregation did not remain in the area in the summer months. We do not know if these killer whales follow gray whales as they migrate northwards, although this seems likely, but our data suggest that they move outside of the study area used for this mark-recapture analysis. This underscores the fact that our abundance estimate refers to the number of killer whales that used the study area *at some time* during the study period, and does not imply that all the animals remained within the area for the entire duration of the study.

Using distance-sampling data collected on three of the same surveys used here (01LT, 02LT, 03LT), it was estimated that only 251 (95% CI = 97-644) transients were present in the same study area at the time of the line-transect surveys in July and August (Zerbini et al. 2007). This instantaneous estimate of density differs in interpretation from our mark-recapture

assumption of the number of distinct animals *using* the area during the full study duration. Although the line-transect estimate is somewhat less precise than the mark-recapture estimate (LT 95% Confidence Intervals = 97-644; MR 95% Probability Intervals = 255-487), it is clear that much of the difference in estimates could be due to animals that were only encountered in the 03FP spring survey of the mark-recapture analysis, and were therefore likely not present in the July-August line transect sampling period. The complimentary inference obtained by comparing these different estimates of abundance will have to be considered when making energetic calculations of the potential impact of killer whale predation on prey populations. Specifically, it is clear that not all the individuals remain in this area at all times, resulting in temporal and spatial variation in the predation pressure on coastal marine mammals.

## Acknowledgements

We are grateful to the many scientists and crew who participated in the survey efforts. Sue Moore and Caroline Gudmundson contributed killer whale photographs, and Alex Zerbini helped with the design of the line-transect surveys. Brian Fadely, Devin Johnson and Nancy Friday provided comments on earlier drafts, and editorial suggestions were made by Gary Duker and James Lee. JWD was supported by a post-doctoral research associateship from the National Research Council and field efforts were supported by NOAA's Steller Sea Lion Research Initiative, with specific funding from the National Marine Mammal Laboratory, the North Pacific Universities Marine Mammal Research Consortium and the Alaska SeaLife Center. Research was conducted under permits 545-1488-03, 782-1510 or 932-1489-05 issued by the National Marine Fisheries Service.

## References

- Baird RW, Stacey PJ (1988) Variation in saddle patch pigmentation in populations of killer whales (*Orcinus orca*) from British Columbia and Washington State. *Canadian Journal of Zoology* 66: 2582-2585
- Baird RW, Whitehead H (2000) Social organisation of mammal-eating killer whales: group stability and dispersal patterns. *Canadian Journal of Zoology* 78: 2096 - 2105
- Barrett-Lennard LG (2000) Population structure and mating patterns of killer whales, *Orcinus orca*, as revealed by DNA analysis. Ph.D. thesis in Zoology. University of British Columbia, Vancouver, Canada.
- Barrett-Lennard LG, Matkin CO, Durban JW, Saulitis EL, Ellifrit DK (In Press.) Gray Whale Predation and Underwater Prey Caching by Transients Killer Whales at Unimak Island, Alaska. *Marine Ecology Progress Series*
- Barrett-Lennard LG, Smith TG, Ellis GM (1996) A cetacean biopsy system using lightweight pneumatic darts, and its effect on the behavior of killer whales. *Marine Mammal Science* 12: 14-27
- Berkhof J, Van Mechelen I, Gelman A (2003) A Bayesian approach to the selection and testing of mixture models. *Statistica Sinica* 13: 423-442
- Bigg MA, Olesiuk PF, Ellis GM, Ford JKB, Balcomb KC (1990) Social Organization and Genealogy of Resident Killer Whales (*Orcinus orca*) in the Coastal Waters of British Columbia and Washington State. In: Hammond P (ed) *Individual Recognition of Cetaceans: Use of Photo-Identification and Other Techniques to Estimate Population Parameters*, Cambridge, pp 383-405
- Black NA, Schulman-Janiger A, Ternullo RL, Guerrero-Ruiz M (1997) Killer whales of California and Western Mexico: A catalog of photo-identified individuals. National Oceanic and Atmospheric Administration Technical Memorandum. NMFS-SWFWC-247. United States Department of Commerce.
- Brooks SP, Gelman A (1998) General methods for monitoring convergence of iterative simulations. *Journal of Computational and Graphical Statistics* 7: 434-455
- Calambokidis J, Barlow J (2004) Abundance of blue and humpback whales in the eastern North Pacific estimated by capture-recapture and line-transect methods. *Marine Mammal Science* 20: 63-85
- Carothers AD (1973a) Capture-recapture methods applied to a population with known parameters. *Journal of Animal Ecology* 42: 125-146
- Carothers AD (1973b) The effects of unequal catchability on Jolly-Seber estimates. *Biometrics* 29: 79-100

- Chao A (2001) An Overview of Closed Capture-Recapture Models. *Journal of Agricultural, Biological and Environmental Statistics*. 6: 158-175
- Coull BA, Agresti AA (1999) The use of mixed logit models to reflect heterogeneity in capture-recapture studies. *Biometrics* 55: 294-301
- Dahlheim ME, Ellifrit DK, Swenson JD (1997) Killer whales of southeast Alaska. A catalogue of photo-identified individuals. Day Moon, Seattle
- Dahlheim ME, Heyning JE (1999) Killer whale *Orcinus orca* (Linnaeus, 1758). In: Ridgway S, Reynolds J (eds) *Handbook of marine mammals*, volume 6. Academic Press, San Diego
- Deecke VB, Ford JKB, Slater PJB (2005) The vocal behaviour of mammal-eating killer whales: communicating with costly calls. *Animal Behaviour* 69: 395-405
- Demaster DP, Trites AW, Clapham P, Mizroch S, Wade P, Small RJ, Ver Hoef J (2006) The sequential megafaunal collapse hypothesis: Testing with existing data. *Progress in Oceanography* 68: 329-342
- Dey D, Muller P, Sinha D (1999) *Practical nonparametric and semiparametric Bayesian statistics*. Springer, New York
- Dorazio RM, Royle AJ (2003) Mixture models for estimating the size of a closed population when capture rates vary among individuals. *Biometrics* 59: 351-364
- Durban JW, Elston DA, Ellifrit DK, Dickson E, Hammond PS, Thompson PM (2005) Multi-site mark-recapture for cetaceans: population estimates with Bayesian model averaging. *Marine Mammal Science* 21: 80-92
- Durban JW, Elston DE (2005) Mark-recapture with occasion and individual effects: abundance estimation through Bayesian model selection in a fixed dimensional parameter space. *Journal of Agricultural, Biological and Environmental Statistics*. 10: 291-305
- Escobar MD, West M (1995) Bayesian density estimation and inference using Mixtures. *Journal of the American Statistical Association* 90: 577-588
- Estes J, Doak D, Springer A, Williams T (2009) Causes and consequences of marine mammal population declines in southwest Alaska: a food-web perspective. *Philosophical Transactions B* 364: 1647
- Estes JA, Tinker MT, Williams TM, Doak DF (1998) Killer whale predation on sea otters linking oceanic and nearshore ecosystems. *Science* 282: 473-476
- Fienberg SE, Johnson MS, Junker BW (1999) Classical multilevel and Bayesian approaches to population size estimation using multiple lists. *Journal of the Royal Statistical Society, Series A* 162: 383-405
- Ford JKB (1989) Acoustic behaviour of resident killer whales (*Orcinus orca*) off Vancouver Island, British Columbia. *Canadian Journal of Zoology* 67: 727-745
- Ford JKB, Ellis GM (1999) *Transients: mammal-hunting killer whales*. University of British Columbia Press, Vancouver, BC.
- Ford JKB, Ellis GM, Balcomb KC (2000) Killer Whales: The natural history and genealogy of *Orcinus orca* in British Columbia and Washington State. UBC Press, Vancouver, B.C.
- Ford JKB, Ellis GM, Barrett-Lennard LG, Morton AB, Palm RS, Balcomb KC (1998) Dietary specialization in two sympatric populations of killer whales (*Orcinus orca*) in coastal British Columbia and adjacent waters. *Canadian Journal of Zoology* 76: 1456-1471
- Fritz LW, Stinchcomb C (2005) Aerial, ship, and land-based surveys of Steller sea lions (*Eumetopias jubatus*) in the western stock in Alaska, June and July 2003 and 2004. . National Oceanic and Atmospheric Administration Technical Memorandum NMFS-AFSC-135. United States Department of Commerce
- Gelfand AE, Ghosh SK (1998) Model choice: A minimum posterior predictive loss approach. *Biometrika* 85: 1-11
- Gelman A, Carlin JB, Stern HS, Rubin DB (1995) *Bayesian Data Analysis*. Chapman and Hall, London, UK
- Gelman A, Meng X, Stern H (1996) Posterior predictive assessment of model fitness via realized discrepancies. *Statistica Sinica* 6: 733-759
- Guenette S, Heymans SJJ, Christensen V, Trites AW (2006) Ecosystem models show combined effects of fishing, predation, competition, and ocean productivity on Steller sea lions (*Eumetopias jubatus*) in Alaska. *Canadian Journal of Fisheries and Aquatic Sciences* 63: 2495-2517
- Herman DP, Burrows DG, Wade PR, Durban JW, LeDuc RG, Matkin CO, Barrett-Lennard L, Krahn MM (2005) Feeding ecology of eastern North Pacific killer whales *Orcinus orca* from fatty acid, stable isotope, and organochlorine analyses of blubber biopsies. *Marine Ecology Progress Series* 302: 275-291
- Hoelzel AR, Dahlheim M, Stern SJ (1998) Low genetic variation among killer whales (*Orcinus orca*) in the eastern north Pacific and genetic differentiation between foraging specialists. *Journal of Heredity* 89: 121-128
- Hoelzel AR, Dover GA (1990) Genetic differentiation between sympatric killer whale populations. *Heredity* 66
- Hoelzel AR, Natoli A, Dahlheim ME, Olavarria C, Baird RW, Black NA (2002) Low worldwide genetic diversity in the killer whale (*Orcinus orca*): implications for demographic history. *Proc. R. Soc. Lond. B* 269: 1467-1473

- Krahn MM, Herman DP, Matkin CO, Durban JW, Barrett-Lennard L, Burrows DG, Dahlheim ME, Black N, LeDuc RG, Wade PR (2007) Use of chemical tracers in assessing the diet and foraging regions of eastern North Pacific killer whales. *Marine Environmental Research* 63: 91-114
- LeDuc R (2004) Report of the results of the 2002 survey for North Pacific right whales. National Oceanic and Atmospheric Administration Technical Memorandum NMFS-SWFSC-357. United States Department of Commerce
- Loughlin TR, Perlov AS, Vladimirov VA (1992) Range-wide survey and estimation of total number of steller sea lions in 1989. *Marine Mammal Science* 8: 220-239
- Lunn DJ, Thomas A, Best N, Spiegelhalter D (2000) WinBUGS - a Bayesian modelling framework: concepts, structure and extensibility. *Statistics and Computing* 10: 325-337
- Lusseau D, Wilson B, Hammond PS, Grellier K, Durban JW, Parsons KM, Barton TR, Thompson PM (2006) Quantifying the influence of sociality on population structure in bottlenose dolphins. *Journal of Animal Ecology* 75: 14-24
- Matkin CO, Barrett-Lennard LG, Yurk H, Ellifrit D, Trites AW (2007) Ecotypic variation and predatory behavior among killer whales (*Orcinus orca*) off the eastern Aleutian Islands, Alaska. *Fishery Bulletin* 105: 74-87
- Matkin CO, Ellis G, Saulitis E, Barrett-Lennard L, Matkin D (1999) Killer whales of southern Alaska. North Gulf Oceanic Society, Homer, Alaska
- Merrick RL, Loughlin TR (1997) Foraging behavior of adult female and young-of-the-year Steller sea lions in Alaskan waters. *Canadian Journal of Zoology* 75: 776-786
- Mizroch SA, Rice DW (2006) Have North Pacific killer whales switched prey species in response to depletion of the great whale populations? *Marine Ecology Progress Series* 310: 235-246
- Neal RM (2000) Markov Chain Sampling Methods for Dirichlet Process Mixture Models. *Journal of Computational and Graphical Statistics* 9: 249-265
- Otis DL, Burnham KP, White GC, Anderson DR (1978) Statistical inference from capture data on closed animal populations. *Wildlife Monographs* 62: 1-135
- Pace ML, Cole JJ, Carpenter SR, Kitchell JF (1999) Trophic cascades revealed in diverse ecosystems. *Trends in Ecology and Evolution* 14: 483-488
- Parsons K, Balcomb K, Ford J, Durban J (2009) The social dynamics of southern resident killer whales and conservation implications for this endangered population. *Animal Behaviour* 77: 963-971
- Pledger S (2000) Unified Maximum Likelihood Estimates for Closed Capture-Recapture Models Using Mixtures. *Biometrics* 56: 434-442
- Saulitis EL, Matkin CO, Barrett-Lennard L, Heise K, Ellis GM (2000) Foraging strategies of sympatric killer whale (*Orcinus orca*) populations in Prince William Sound, Alaska. *Marine Mammal Science* 16: 94-109
- Sease JL, Gudmundson C (2002) Aerial and land-based surveys of Steller sea lions (*Eumetopias jubatus*) from the western stock in Alaska, June and July 2001 and 2002. National Oceanic and Atmospheric Administration Technical Memorandum NMFS-AFSC-131. United States Department of Commerce
- Seber GAF (1982) The estimation of animal abundance and related parameters. Macmillan Publishing Company, New York, NY
- Sinclair EH, Moore SE, Friday NA, Zeppelin TA, Waite JM (2005) Do patterns of Steller sea lion (*Eumetopias jubatus*) diet, population trend and cetacean occurrence reflect oceanographic domains from the Alaska Peninsula to the central Aleutian Islands? *Fisheries Oceanography* 14: 223-242
- Springer A, Estes J, Van Vliet G, Williams T, Doak D, Danner E, Forney K, Pfister B (2003) Sequential megafaunal collapse in the North Pacific Ocean: An ongoing legacy of industrial whaling? *Proceedings of the National Academy of Sciences* 100: 12223-12228
- Springer AM, Estes JA, van Vliet GB, Williams TM, Doak DF, Danner EM, Pfister B (2008) Mammal-eating killer whales, industrial whaling, and the sequential megafaunal collapse in the North Pacific Ocean: A reply to critics of Springer et al. 2003. *Marine Mammal Science* 24: 414-442
- Stevens TA, Duffield DA, Asper ED, Hewlett KG, Bolz A, Gage LJ, Bossart GD (1989) Preliminary findings of restriction fragment differences in mitochondrial DNA among killer whales (*Orcinus orca*). *Canadian Journal of Zoology* 67: 2592-2595
- Stevick PT, Allen J, Clapham PJ, Friday N, Katona SK, Larsen F, Lien J, Mattila DK, Palsboll PJ, Sigurjonsson J, Smith TD, Oien N, Hammond PS (2003) North Atlantic humpback whale abundance and rate of increase four decades after protection from whaling. *Marine Ecology Progress Series* 258: 263-273
- Towell RG, Ream RR, York AE (2006) Decline in northern fur seal (*Callorhinus ursinus*) pup production on the Pribilof Islands. *Marine Mammal Science* 22: 486-491
- Trites AW, Deecke VB, Gregr EJ, Ford JKB, Olesiuk PF (2007) Killer whales, whaling, and sequential megafaunal collapse in the North Pacific: a comparative analysis of the dynamics of marine mammals in Alaska and British Columbia following commercial whaling. *Marine Mammal Science* 23: 751-765

- Wade PR, Burkanov VN, Dahlheim ME, Friday NA, Fritz LW, Loughlin TR, Mizroch SA, Muto MM, Rice DW, Barrett-Lennard LG, Black NA, Burdin AM, Calambokidis J, Cerchio S, Ford JKB, Jacobsen JK, Matkin CO, Matkin DR, Mehta AV, Small RJ, Straley JM, McCluskey SM, VanBlaricom GR (2007) Killer whales and marine mammal trends in the North Pacific - a re-examination of evidence for sequential megafauna collapse and the prey-switching hypothesis. *Marine Mammal Science* 23: 766-802
- Waite JM, Friday NA, Moore SE (2002) Killer whale (*Orcinus orca*) distribution and abundance in the Central and Southeastern Bering Sea, July 1999 and June 2000. *Marine Mammal Science* 18: 779-786
- West M (1992) Mixture models, Monte Carlo, Bayesian updating and dynamic models. *Journal of Statistical Planning and Inference*. 62: 325-333
- Williams TM, Estes JA, Doak DF, Springer AM (2004) Killer appetites: assessing the role of predators in ecological communities. *Ecology* 85: 3373-3384
- Wilson B, Hammond PS, Thompson PM (1999) Estimating Size and Assessing Trends in a Coastal Bottlenose Dolphin Population. *Ecological Applications* 9: 288-300
- Zerbini AN, Waite JM, Durban JW, Dahlheim ME, Wade PR (2007) Estimating abundance of killer whales in the nearshore waters of the Gulf of Alaska and Aleutian Islands using line transect sampling. *Marine Biology* 150: 1033-1045

**Table 1.** Summary data for each of the 14 surveys from which photo-identification data were collected. The strata with encounters are spatially referenced in Figure 1. Only high quality images were used, where Q2 indicated a usable image of the saddle region only, Q3 indicated a usable fin only and Q4 indicated both the fin and saddle in useable quality. Primary marks consisted of notches in the dorsal fin that could be used alone to uniquely identify individuals. Secondary makings included oval scars, saddle patch pigmentation patterns, linear scars on the saddle patch, and variation in dorsal fin shape, and were used in combination to identify individuals if primary marks did not exist.

Survey	Strata with encounters	# Encounters	# Photographs	Images				Individuals	
				Total	Q2	Q3	Q4	Primary marks	Secondary marks
01LT	2, 9, 10, 11, 12, 14	7	779	928	3	62	127	8	7
01KK	1	5	78	101	1	22	22	3	3
01EA	11	1	48	54	3	12	24	1	2
02AH	11, 12	4	305	404	7	77	67	17	7
02LT	12, 14	3	370	489	7	95	68	8	7
02MA	3, 11, 12	3	52	57	1	6	1	2	0
02MF	11	1	52	95	0	5	2	5	0
02KK	1	8	430	592	15	112	94	6	2
02EA	11, 12	4	368	544	13	104	184	8	8
02TX	11	1	57	113	4	11	20	4	4
03FP	10, 12	11	1503	1928	43	442	471	38	21
03LT	10, 11, 12, 16	4	307	436	5	129	62	11	2
03KK	1	7	189	252	3	47	43	5	4
03EA	11	4	394	496	5	99	217	13	7
<b>TOTAL</b>	<b>1,2,3,9,10,11,12,14,16</b>	<b>63</b>	<b>4932</b>	<b>6489</b>	<b>110</b>	<b>1223</b>	<b>1402</b>	<b>129</b>	<b>74</b>

**Table 2.** The estimated mean square predicted error (MSPE) inferring the fit of each model to the photo-id data, along with the posterior median (95%probability interval) estimates for the number of clusters of individuals from the three models with mixture components and the abundance of distinctively marked individuals,  $N$ , for all models.

<i>Model</i>	$\log\{p_{ij}/(1-p_{ij})\}$	<i>MSPE</i>	<i>Clusters</i>	<i>N</i>
$M_0$	$\mu$	0.17	-	333 (272 - 426)
$M_\alpha$	$\mu + \alpha_i$	0.16	-	315 (257 - 400)
$M_\theta$	$\mu + \theta_j$	0.17	1 (1 - 2)	335 (271 - 433)
$M_{\alpha\theta}$	$\mu + \alpha_i + \theta_j$	0.16	1 (1 - 2)	317 (258 - 411)
$M_{\alpha\theta\lambda}$	$\mu + \alpha_i + \theta_j + \lambda_{ij}$	0.05	7 (7 - 10)	275 (210 - 372)

**Table 3.** The off-diagonal elements of the covariance matrix  $\Sigma$ , showing the associations between pairs of surveys in terms of their similarity in coverage of specific clusters. The posterior median of each covariance is displayed, and denoted in bold if the 95% probability intervals of the distribution did not encompass zero (no association).

	01LT	01KK	01EA	02AH	02LT	02MA	02MF	02KK	02EA	02TX	03FP	03LT	03KK	03EA
01LT		-8	4	3	<b>-15</b>	5	-4	<b>-11</b>	11	8	15	11	<b>-13</b>	-3
01KK			<b>-15</b>	-1	-8	-4	-22	<b>15</b>	<b>-22</b>	<b>-26</b>	35	<b>-27</b>	<b>19</b>	<b>-37</b>
01EA				-8	<b>14</b>	2	<b>35</b>	<b>-13</b>	<b>23</b>	<b>33</b>	<b>-51</b>	<b>36</b>	<b>-17</b>	<b>51</b>
02AH					-8	1	<b>-37</b>	-4	-4	<b>-24</b>	28	<b>-29</b>	-3	<b>-36</b>
02LT						-2	<b>40</b>	-3	13	<b>26</b>	<b>-78</b>	<b>24</b>	-4	<b>59</b>
02MA							1	-5	4	4	-3	4	-5	4
02MF								-11	<b>39</b>	<b>83</b>	<b>-125</b>	<b>87</b>	-19	<b>127</b>
02KK									<b>-20</b>	<b>-19</b>	23	<b>-22</b>	<b>20</b>	-25
02EA										<b>41</b>	<b>-58</b>	<b>45</b>	<b>-27</b>	<b>62</b>
02TX											<b>-99</b>	<b>77</b>	<b>-27</b>	<b>105</b>
03FP												<b>-100</b>	34	<b>-178</b>
03LT													<b>-31</b>	<b>111</b>
03KK														<b>-39</b>
03EA														

**Figure 1.** Map of the study area, defined by rectangular strata (light shading, numbered 1-16) covering the coastal waters of the Gulf of Alaska and Aleutian Islands. Each data point (solid circle) represents the location of a single encounter with a group of mammal-eating killer whales.

**Figure 2.** Photographs showing examples of natural markings used to identify individual killer whales. Notches in the dorsal fin (a) were considered primary markings that could be used to uniquely identify individuals. Secondary markings such as oval scars (a and b), saddle patch pigmentation patterns (c), linear scars on the saddle patch (c), and variation in dorsal fin shape (a,b,c) were used in combination to identify individuals. Some individuals were not deemed to be sufficiently distinctive for long-term matches (d), but subtle secondary markings allowed them to be distinguished within the group.

**Figure 3.** Cumulative individuals plotted against cumulative identifications, with data points (filled squares) presented for each survey in chronological sequence. The hypothetical 1:1 discovery rate (solid diagonal) is plotted for reference, and the three calendar years (2001, 2002, 2003) are separated by broken lines.

**Figure 4.** Posterior probability distribution (bars) and prior probability distribution (broken line) for the overall abundance,  $P$ , of killer whales, estimated using the best-fitting  $M_{\alpha\theta_i}$  model.







