Statistical Methodology 17 (2014) 46–66

Contents lists available at ScienceDirect

Statistical Methodology

journal homepage: www.elsevier.com/locate/stamet

A spatial hierarchical model for abundance of three ice-associated seal species in the eastern Bering Sea

Jay M. Ver Hoef*, Michael F. Cameron, Peter L. Boveng, Josh M. London, Erin E. Moreland

National Marine Mammal Laboratory, Alaska Fisheries Science Center, NOAA National Marine Fisheries Service, 7600 Sand Point Way NE, Bldg 4, Seattle, WA 98115-6349, USA

ARTICLE INFO

Article history:
Received 8 June 2012
Received in revised form
6 December 2012
Accepted 17 March 2013

Keywords:
Conditional autoregressive
Line transect
Sampling
Spatial linear model
Detection

ABSTRACT

Estimating the abundance of seals inhabiting sea ice is complicated because the areas are large, the ice area and distribution may change rapidly, and it is impractical to detect and count a (typically unknown) portion of the population that is in the water, rather than hauled out on the ice. We propose a method for resolving these issues by using a series of daily estimates that are imprecise by themselves, but yield an acceptable estimate when they are combined. Population surveys of spotted seals, ribbon seals and bearded seals were conducted over 279,880 km^2 of the Bering Sea between 13 April and 26 May 2007. During that period, the sea-ice conditions and spatial distribution of seals changed dramatically. We surveyed 2748 km^2 using line transect methods from a helicopter deployed from the US Coast Guard icebreaker *Healy*. Corrections for incomplete availability of seals used a generalized linear mixed model for seal haul-out probability using sensors on seals with data transmitted by satellite. We accounted for incomplete detection through standard distance-sampling methods along with a double-observer model. The availability and detection models were combined in a data model for local abundance in space and time. To accommodate the shifting ice and seal movements, we then developed a hierarchical spatially-autocorrelated regression model using remotely sensed sea ice concentration data to predict

* Correspondence to: University of Alaska International Arctic Research Center, P.O. Box 757345, Fairbanks, AK 99775-7345, USA. Tel.: +1 907 456 1995; fax: +1 907 458 3737.
E-mail address: jay.verhoef@noaa.gov (J.M. Ver Hoef).

1572-3127/$ – see front matter. Published by Elsevier B.V.
http://dx.doi.org/10.1016/j.stamet.2013.03.001

abundance at each survey date. While abundance estimation was very imprecise for each date, we were able to combine them to obtain good estimates of overall population abundance even though the population was spatially dynamic. The proposed hierarchical model combined submodels and accounted for their sources of uncertainty. Spotted seals were most abundant within the study area (233,700, 95% CI 137,300–793,100), followed by bearded seals (61,800, 95% CI 34,900–171,600) and ribbon seals (61,100, 95% CI 35,200–189,300).

Published by Elsevier B.V.

1. Introduction

The Bering Sea is an area of dynamic seasonal changes in sea-ice cover. This sea ice provides important habitat for several marine mammals, including seals, walrus, and polar bears. Sea ice is also one of the environmental features most vulnerable to global climate warming, possibly threatening the animals that depend on it. To begin to monitor population changes, we conducted surveys of spotted seals (*Phoca largha*), ribbon seals (*Histriophoca fasciata*) and bearded seals (*Erignathus barbatus*) on ice within an area of 81,600 km^2 in the Bering Sea (Fig. 1) between 13 April and 10 June 2007. All three seal species have been or are being considered for listing as threatened or endangered species under the US Endangered Species Act. Our goal in this article is to develop abundance estimates for all three seal species in an important portion of their ranges.

The analysis of the survey data posed several problems. For an observational model, there were layered probabilities for encountering seals. First, not all seals could be seen at the time of surveys, as some were in the water where they are inconspicuous. We addressed this by using a model for the probability of each seal species being hauled out on the ice. Haul-out data were obtained from satellite-linked dive recorders (SDRs) fixed to a sample of seals for each species. However, given that a seal is on the ice and available to be seen, it may yet be missed. This was addressed by using line transect surveying methods [9], recording distances to seals when observed from helicopter transects over ice.

The second main problem was due to the relatively long time frame, approximately two months, in which surveys occurred. Each seal species tended to be associated with certain ice conditions, and the ice conditions changed dramatically during the survey period, shrinking by millions of hectares. This caused the population distribution to change dramatically during the survey period. Thus, it was not sensible to estimate abundance using traditional line transect sampling theory that assumes transects are replicate samples for a fixed population because whole parts of a population may be unknowingly resampled at a later date. Maps of the area with ice extent at various dates and the locations of transects are shown in Fig. 1.

1.1. Model outline

This article addressed the shifting ice distribution and the incomplete observation of seals by creating a hierarchical modeling framework, described in Sections 2 and 3. The first stage of the hierarchical model depended upon a "sampling" or "data" model. In Section 2, we built the data model where there were two sources of data: (1) aerial surveys, which used distance-sampling methods, and (2) satellite transmitters, which used a temporally-autocorrelated logistic regression model. The vector of estimated detection probabilities from distance sampling was denoted as $\hat{\mathbf{p}}_d$, and the estimated probability that a seal was out of the water was denoted as $\hat{\mathbf{p}}_h$, so the combined probability of seeing the observed seals was $\hat{\mathbf{p}} = \hat{\mathbf{p}}_d \odot \hat{\mathbf{p}}_h$. Note that in general, throughout the manuscript, we use the hat symbol, $\hat{\mathbf{p}}$, to denote an estimate. A spatial grid of square cells was created in the Bering sea, and the estimated probabilities of observed seals were used to estimate density per grid cell per date by $\hat{D}_{t,i} = \frac{1}{a_{t,i}} \sum_{j=1}^{n_{t,i}} \frac{s_{t,i,j}}{\hat{p}_{t,i,j}}$ for the ith cell on the tth date, where $\hat{p}_{t,i,j} \in \hat{\mathbf{p}}$, $a_{t,i}$ is the area surveyed

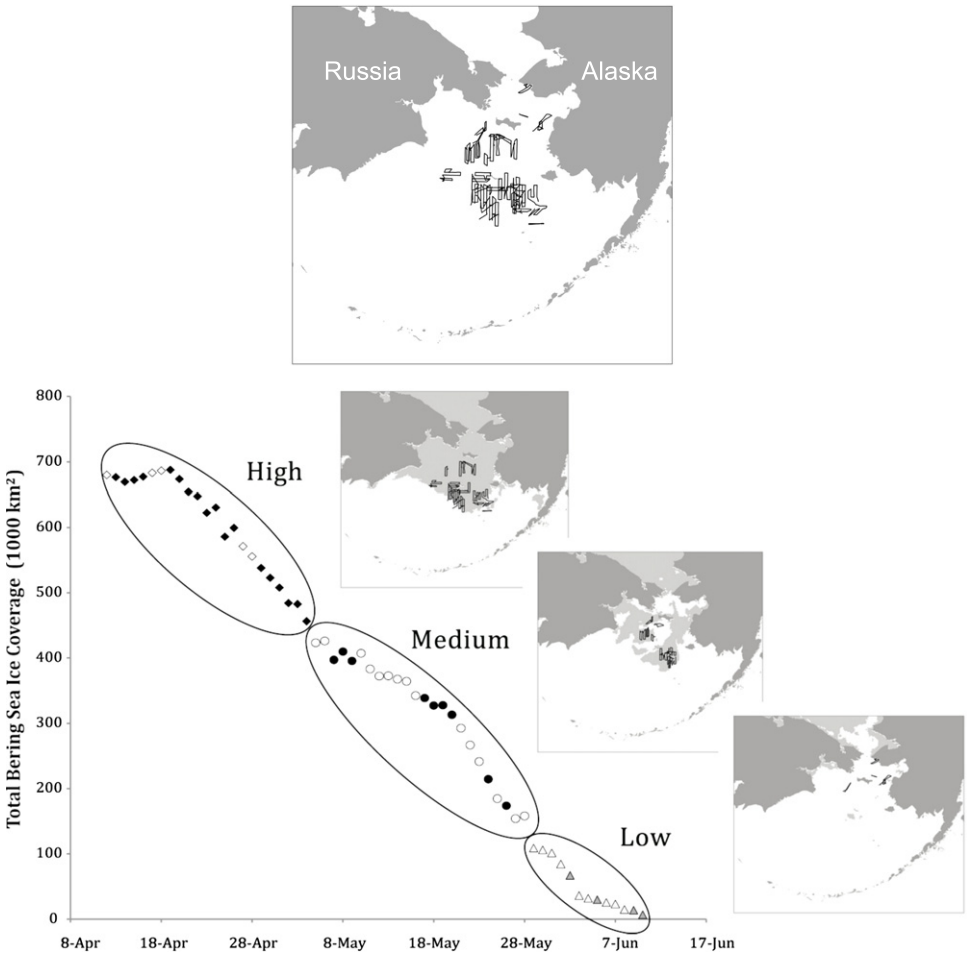

Fig. 1. Top: line transect surveys in the Bering Sea, occurring between 8 April and 10 June 2007. Bottom: ice extent in the Bering Sea, shown as the medium shade gray and defined as ice concentration greater than 10% ice, between 8 April and 10 June 2007. Three representative dates are shown for each time period that we arbitrarily labeled as "High", "Medium", and "Low". All transects are shown on top of the ice extent within each time period. All symbols indicate days with images for computing ice extent, and solid symbols are days when transects were flown.

in the ith cell on the tth date, and $s_{t,i,j}$ is the group size (number of seals) for the jth observation, $j = 1, 2, \ldots, n_{t,i}$, per t, i combination. Let $\mathbb{O} = \{t, i\}$ be the set of grid-cell/days in the grid $\times$ survey days that contained survey flights, and let $\mathbb{S} = \{t, i\}$ be the set of cell-days in the grid $\times$ survey days that contained observed seals ($\mathbb{S} \subset \mathbb{O}$). We denote the set of all observed $\{\hat{D}_{\{t,i\}\in\mathbb{S}}\} = \hat{\mathbf{d}}_o$, where we also take care to estimate $\text{var}(\hat{\mathbf{d}}_o) = \hat{\mathbf{C}}_o$. There were many cells that had zero counts during flights, so we denote the binary vector $\mathbf{b}_o$ for the observed counts, where $\{B_{\{t,i\}\in\mathbb{O}}\} = \mathbf{b}_o$ is one if the count was greater than zero, and zero otherwise. Note that $\mathbf{b}_o$ has length greater than $\hat{\mathbf{d}}_o$.

In Section 3, we build the hierarchical model following the general frame work set out by Berliner [3], [process, parameters|data] $\propto$ [data|process] [process|parameters] [parameters], where the brackets "$[\cdots]$" refer to a distribution and the vertical bar "$\cdots|\cdots$" denotes conditional dependence; also see [14]. We take part of the data model to be $[\log(\hat{\mathbf{d}}_o)|\mathbf{m}_o, \text{diag}(1/\hat{\mathbf{d}}_o)\hat{\mathbf{C}}_o\text{diag}(1/\hat{\mathbf{d}}_o)]$; a lognormal distribution with mean $\mathbf{m}_o$ and covariance matrix $\text{diag}(1/\hat{\mathbf{d}}_o)\hat{\mathbf{C}}_o\text{diag}(1/\hat{\mathbf{d}}_o)$ (using the

delta method for var[log($\hat{\mathbf{d}}_o$)]). Then $\mathbf{m}_o$ is modeled with a spatial regression model, $[M_{t,i}|X_{t,i}, \boldsymbol{\omega}, W_{t,i}]$, where $X_{t,i}$ is an ice concentration covariate for the ith cell on date t taken from satellite images, with regression parameters $\boldsymbol{\omega}$, and spatial random effects $\{W_{t,i}\}$ that have a spatial conditional-autoregressive model (CAR) [4] among the i cells for the tth survey. The spatial random effects form a random field, and random fields are assumed independent for each t. Denote $\mathbf{m}$ as an extension of $\mathbf{m}_o$ to include unobserved cells. During the Markov chain Monte Carlo (MCMC) sampling, all $\mathbf{m}$ are treated as unobserved latent variables, which essentially allows spatial smoothing (with covariates) for observed cells, and spatial prediction into unobserved cells. We then model the spatial binary process $\mathbf{b}_o$ as a Bernoulli process, $[B_{t,i}|Q_{t,i}, \chi_{t,i}]$, where $Q_{t,i}$ is the true presence/absence of a seal, which may not have been observed because the whole grid cell was not surveyed. The probability of sighting is modeled with a Bernoulli process $[\chi_{t,i}|\boldsymbol{\gamma}, a_{t,i}]$ with regression parameters $\boldsymbol{\gamma}$ that scale the surveyed area $a_{t,i}$. Then the true presence/absence $Q_{t,i}$ is modeled as Bernoulli $[Q_{t,i}|X_{t,i}, \boldsymbol{\kappa}, Z_{t,i}]$, where again $X_{t,i}$ is an ice concentration covariate, this time with regression parameters $\boldsymbol{\kappa}$, and spatial random effects $\{Z_{t,i}\}$ that have a spatial conditional-autoregressive model (CAR); here $\pi_{t,i} = E(Q_{t,i}) = f(X_{t,i}, \boldsymbol{\kappa}, W_{t,i})$. Again, the spatial random effects $\{Z_{t,i}\}$ form a random field, and random fields are assumed independent for each t. Denote $\boldsymbol{\pi}$ to include all unobserved cells for all surveys. During the Markov chain Monte Carlo (MCMC) sampling, all $\boldsymbol{\pi}$ are treated as unobserved latent variables, which essentially allows spatial smoothing (with covariates) for observed cells, and spatial prediction into unobserved cells.

The modeling in Sections 2 and 3 allowed us to solve both of our main problems. Using the posterior distributions of $\mathbf{m}$ and $\boldsymbol{\pi}$, we formed abundance estimates as $\hat{Y}_{t,\bullet} = \sum_{i=1}^{N_A} \exp(M_{t,i})\pi_{t,i}A_i$, where A_i is the total area of the ith grid cell and $N_A = 192$ is the number of spatial grid cells. These were very imprecise estimates, yet they did not suffer from the possibility of seal movements through time. We then examined these estimates for any temporal trends; if none existed, then we felt justified in combining them into an overall estimate, $\hat{Y}_{\bullet,\bullet} = \frac{1}{N_T}\sum_{t=1}^{N_T}\sum_{i=1}^{N_A} \exp(M_{t,i})\pi_{t,i}A_i$, that was much more precise, where $N_T = 26$ was the number of daily surveys.

We summarize and discuss our approach in Section 4. Appendix A gives details on fitting the detection function, and Appendix B provides details on a multivariate extension to the variance and covariances of direct products [16,5].

2. The data model

Following the terminology of [3] we begin with the data model. The data model is a statistical model for how the data were observed/obtained. It was developed from two primary parts: a distance sampling model from line transects collected from helicopter surveys, and a haul-out model from SDRs fixed to an independent sample of seals.

2.1. Helicopter survey data

In the spring of 2007, researchers from the Alaska Fisheries Science Center's National Marine Mammal Laboratory, which is part of the National Oceanic and Atmospheric Administration's (NOAA) National Marine Fisheries Service (NMFS), conducted aerial surveys for ribbon, bearded, and spotted seals in the US sector of the Bering Sea [12]. The surveys were conducted from a Bell Long Ranger[1] helicopter based aboard the US Coast Guard icebreaker *Healy*. The cruises ranged throughout the pack ice of the eastern and central Bering Sea, providing access to areas not surveyed since the 1970s and 1980s. Note that ringed seals (*Phoca hispida*) were also found in the sea ice of our study area, but we did not include them in our analysis. The number of ringed seal sightings was low, possibly due to a combination of a preponderance of survey effort away from near-shore areas favored by ringed seals and a greater tendency for ringed seals to be disturbed into the water by the helicopter, and therefore to be missed by observers.

[1] Reference to trade names does not imply endorsement by the National Marine Fisheries Service, NOAA.

We conducted line transect surveys from the helicopter whenever the *Healy* was near ice and when weather conditions were conducive to flying between approximately 09:00 and 16:00 (local apparent time), which corresponds to the timing of peak seal haul-out probability. By local apparent time, we mean that for each 15° of latitude west of Greenwich Mean Time (GMT), one hour was subtracted; thus, we used GMT minus 11 h for Bering Sea observations which puts the sun overhead at approximately noon at these coordinates. Each flight had two to three observers and was flown at a target altitude of 118 m (400 ft) at speeds of 80–95 knots. Only seals hauled out on ice were observed and recorded. The distance from each seal to the helicopter's track line was calculated using a sighting bar mounted on each observer's window. In all, 1567 seals were observed during approximately 49 h of survey effort covering 8175 km of survey line on 44 flights. Because ice conditions decayed markedly toward the end, our analysis used only the first 27 flights (the high and medium ice extent in Fig. 1).

We have both mark–recapture and distance data from the helicopter surveys, and when combined they form mark–recapture distance sampling (MRDS) models. We will follow the assumptions of point independence [19,6]. In the point independence models, the likelihoods are independent for estimation of a detection function and estimation of probability of detection at a distance zero, so we model them separately in the next few sections, and then combine them.

2.2. Modeling detection from distance data

Groups of seals ($\geq$1 individual) were classified into distance bins. A clear acrylic strip was fixed to the window frame several centimeters inside from the helicopter window, and by aligning marks on the strip with corresponding marks on the window, the observer could consistently classify observations into distance bins. It was not possible to see directly beneath the helicopter, so the first cutoff value, set to 0, was actually about 85.4 m from the track line of the helicopter, and defined by an angle of 55° from horizontal. Bin cutoff angles were set at 55°, 50°, 40°, 30°, 20°, 10°, and 0° from horizontal. The helicopter flew at a target altitude of 122 m (400 ft) that varied $\pm$15 m. This created distance bin cutoffs of $\mathcal{F} = \{0, 16.9, 59.9, 125.8, 249.6, 606.3, \infty\}$, where recall that 0 is actually offset by 85.4 m from directly beneath the helicopter. The seven cutoff values create six distance classes that we denote $\mathcal{D} = 1, 2, \ldots, 6$, where $\mathcal{D} = 1$ is the distance class $0 - 16.9$, $\mathcal{D} = 2$ is the distance class $16.9 - 59.9$, etc.

We assumed an exponential power series model [22] for detection, which has a shape parameter that allows it more flexibility than the half-normal, which is often used in distance sampling methods. We modeled detection, conditional on distance x, as

$$g(x) = \exp\{-(x/\sigma)^{1/\lambda}\}. \tag{1}$$

We allowed covariates to enter through σ; that is, for the jth observation of species k, $g_{k,j}(x)$, we let

$$\log(\sigma_{k,j}) = \beta_0 + \tau_k + \beta_1 s_{k,j}, \tag{2}$$

where β_0 was an overall intercept, τ_k was a categorical effect for species k, $s_{k,j}$ was the number of observed seals, and β_1 was a regression coefficient. Assuming a spatially uniformly placed location, given that a seal group was detected, the conditional probability that the jth observation of species k was in the $\mathcal{D}$th distance class was

$$p(\mathcal{D}_{k,j}|\boldsymbol{\theta}) = \frac{\int_{l_{\mathcal{D}_j}}^{u_{\mathcal{D}_j}} g_{k,j}(x|\boldsymbol{\theta})dx}{\int_0^\infty g_{k,j}(x|\boldsymbol{\theta})dx}, \tag{3}$$

where $\boldsymbol{\theta}$ is the collection of all parameters: $\boldsymbol{\theta} = (\beta_0, \beta_1, \tau_2, \tau_3, \lambda)'$ (we set $\tau_1 = 0$ for identifiability), and where $l_{\mathcal{D}_j}$ and $u_{\mathcal{D}_j}$ are the bin cutoff distances in $\mathcal{F}$; e.g. $l_2 = 16.9$ and $u_2 = 59.9$. We evaluated (3) numerically. Details of fitting the model are contained in Appendix A. The fitted models are shown in Fig. 2, and the fitted detection values at all observed locations are denoted $\hat{\mathbf{p}}_v$ with an estimated covariance matrix of $\text{var}(\hat{\mathbf{p}}_v) \equiv \hat{\mathbf{C}}_v$.

The fitted values $\hat{\mathbf{p}}_v$ need to be adjusted because $g(0) = 1$ in (1), but we did not have perfect detection when distance was 0; i.e, we believe that $g(0) < 1$. Classical distance sampling

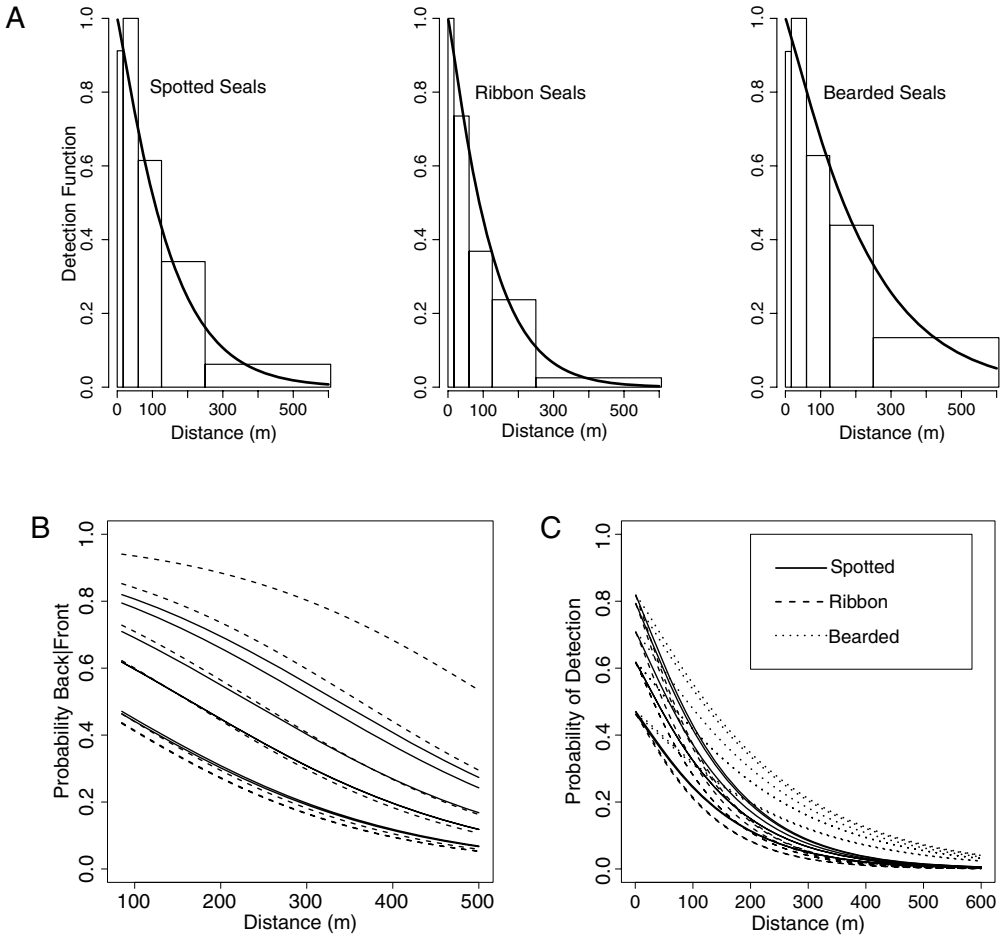

Fig. 2. A. Observed proportions and fitted models for detection functions for each of the three ice seal species. B. Fitted models to each observer from the double-observer data. The solid lines are curves for individual observers using a random effect for each observer. The dashed lines are curves using fixed effects for each observer. C. Final detection models, as a function of distance, for each observer for three species of ice seals.

methods [9], through sampling design and effort, makes the additional assumption that $g(0) = 1$ to get unconditional detection probabilities from the detection function, which allows for abundance estimation. However, more recent developments in distance sampling use double observers to estimate the proper adjustment when $g(0) < 1$ (e.g., [23,7]), which we adopt in the next section.

2.3. Modeling detection from double-observer data

Under the assumption of point independence [19,6] and measuring exact distances, the likelihoods for estimation of conditional detection probability and estimation of absolute detection at distance zero are independent. While the use of bins no longer guarantees this independence, recent research indicates it is a reasonable assumption. Here we describe the analysis of the double-observer data to estimate detection probability at distance zero. In the left front seat of the helicopter we positioned an extra observer on many of the flights. The observations from the left front seat observer were used in a simple way; whenever this observer saw a seal it provided a "trial" for the observer in the back seat on the left side. These binary data ($y_j = 1$ if backseat saw the same seal as front seat, and $y_j = 0$

if backseat observer missed it) were modeled using a generalized linear model with a logit link and Bernoulli distribution; $Y_j \sim \text{Bern}(\mu_j)$. The expected value $\mu_j = f^{-1}(\eta_j) = \frac{e^{\eta_j}}{1+e^{\eta_j}}$ for the jth observation was modeled as

$$\eta_j = \alpha_0 + \alpha_1 c_j + R_j + K_j, \tag{4}$$

where c_j is the center of the distance class for the jth observed seal from the front seat, R_j is a random effect for the different backseat observers, and K_j is a random effect for the species of the jth seal. We fit the model using the lme4 function [2] in R [24]. The random effects K_j due to species were very small and not significant, so we deleted that term from the model. We also tried R_j as a fixed effect. The differences are shown in Fig. 2. There is clearly shrinkage when using the random effects, which seems desirable in this case. That is, the range seen in the fits using R_j as a fixed effect, from approximately 45% to 95% at the closest distances, is much wider than we think is reasonable. Also, the highest curve had a small sample size, so the idea of borrowing strength from the ensemble is preferable; indeed, the highest fixed effect curve is not the highest curve when R_j is treated as a random effect. In what follows, we use the model where observer is a random effect.

Using the lme4 function with logit link and Bernoulli distribution, we obtained estimates of the fixed effects $\hat{\boldsymbol{\alpha}} = (\hat{\alpha_0}, \hat{\alpha_1})$ and random effects $\hat{\mathbf{r}} = (\hat{R}_1, \hat{R}_2, \ldots)$, along with the estimated variance of the random effects $\hat{\text{var}}(R_j) = \hat{\psi}^2$. The covariance matrix of the fitted values for all observations used mixed model theory on variances of fitted fixed and random effects (e.g., [20]). Let $\boldsymbol{\mu} = f^{-1}(\boldsymbol{\eta})$ where the jth element of $\boldsymbol{\eta}$ is η_j and $f^{-1}(\eta_j) = \exp(\eta_j)/(1 + \exp(\eta_j))$; i.e, $f(\mu_j)$ is the logit function. Let $\boldsymbol{\Delta}$ be a diagonal matrix with diagonal elements $\partial\boldsymbol{\mu}/\partial\boldsymbol{\eta}$ and let $\mathbf{A}$ be a diagonal matrix with the variance, in terms of the fit, on the diagonal; i.e., $\mathbf{A}_{i,i} = \exp(\eta_i)/(1 + \exp(\eta_i))^2$. Then, on the logit scale, the variance of an observation is approximated by $\hat{\mathbf{V}} = \hat{\boldsymbol{\Delta}}^{-1}\hat{\mathbf{A}}\hat{\boldsymbol{\Delta}}^{-1}$ using fitted values in $\boldsymbol{\Delta}$ and $\mathbf{A}$. Let $\hat{\mathbf{G}} = \hat{\psi}^2\mathbf{I}$ be the covariance matrix of the random effects. Then the joint covariance matrix of the estimated fixed effects $\hat{\boldsymbol{\alpha}}$ and the estimated random effects $\hat{\mathbf{r}}$ is

$$\hat{\mathbf{C}}_{\alpha,\mathbf{r}} = \begin{pmatrix} \mathbf{X}'\hat{\mathbf{V}}^{-1}\mathbf{X} & \mathbf{X}'\hat{\mathbf{V}}^{-1}\mathbf{Z} \\ \mathbf{Z}'\hat{\mathbf{V}}^{-1}\mathbf{X} & \mathbf{Z}'\hat{\mathbf{V}}^{-1}\mathbf{Z} + \hat{\mathbf{G}}^{-1} \end{pmatrix}^{-1},$$

where $\mathbf{X}$ is the design matrix for the fixed effects in (4) and $\mathbf{Z}$ is the design matrix for the random effects in (4).

Note that Section 2.2 modeled detection decay as a function of distance. Thus, here we are only interested in estimating detection at distance equal to zero as recommended by Laake and Borchers [19] and Borchers et al. [6]. The matrix of prediction coefficients $\mathbf{L}$, where there is a row for each observation, was used to estimate this probability for each observation, which varies only by species and observer. For example, suppose that the first observation was by observer one, then the first row of $\mathbf{L}$ would be $(1, 0, 1, 0, \ldots)$, where the second zero indicates a distance of zero. Then the fitted prediction on the logit scale is

$$\hat{\boldsymbol{\phi}} = \mathbf{L}\begin{pmatrix} \hat{\boldsymbol{\alpha}} \\ \hat{\mathbf{r}} \end{pmatrix} \quad \text{and} \quad \hat{\text{var}}(\hat{\boldsymbol{\phi}}) = \mathbf{L}\hat{\mathbf{C}}_{\alpha,\mathbf{r}}\mathbf{L}' \equiv \hat{\mathbf{C}}_{\phi}.$$

The predictions, transformed back to the probability scale are

$$\hat{\mathbf{p}}_b = \exp(\hat{\boldsymbol{\phi}})/(1 + \exp(\hat{\boldsymbol{\phi}})),$$

with an estimated covariance matrix of

$$\hat{\mathbf{C}}_b = \hat{\mathbf{D}}_b\hat{\mathbf{C}}_{\phi}\hat{\mathbf{D}}_b$$

by using the delta method [15,25], where the ith diagonal element of $\hat{\mathbf{D}}_b$ is $\exp(\hat{\phi}_i)/(1 + \exp(\hat{\phi}_i))^2$.

2.4. Combined MRDS model

Under the point independence model, the estimated probability of detection for the MRDS model is the product of the probability of detection at distance zero times the detection function [6]. The

fitted curves for each observer and species combination, for a single seal, is given in Fig. 2. The fitted detection probability for each observation is

$$\hat{\mathbf{p}}_d = \hat{\mathbf{p}}_b \odot \hat{\mathbf{p}}_v,$$

where $\odot$ is element-wise multiplication, with an estimated covariance matrix of

$$\text{var}(\hat{\mathbf{p}}_d) = \hat{\mathbf{p}}_b\hat{\mathbf{p}}_b' \odot \hat{\mathbf{C}}_v + \hat{\mathbf{p}}_v\hat{\mathbf{p}}_v' \odot \hat{\mathbf{C}}_b - \hat{\mathbf{C}}_v \odot \hat{\mathbf{C}}_b$$

from Appendix B. Denote $\hat{\mathbf{p}}_{d,k}$ as the subset of $\hat{\mathbf{p}}_d$ for the kth species, and let $\hat{\mathbf{C}}_{d,k}$ be comprised of the corresponding rows and columns from var($\hat{\mathbf{p}}_d$). These fitted detection probabilities only apply to seals that were on ice. Some seals were in the water, so further adjustment to detection probabilities is required to account for seals in the water, using haul-out models that are described next.

2.5. Haul-out model

All haul-out data came from 57 spotted seals, 29 bearded seals and 92 ribbon seals. The animals were free-ranging in the Bering Sea during the seasonal period of the surveys (though not all in the same year) affixed with one or two types of satellite data recorders (SDR) (models SPLASH and SPOT5, Wildlife Computers, Redmond, WA). The SPLASH model SDRs were glued to the hair of the animals on the mid-dorsal region or head with five-minute epoxy. The SPOT5 model SDRs were attached to the inter-digital webbing of one of the hindflippers. All tags were programmed to record an electrical resistance value of the surrounding medium (salt-water or air/fresh-water). For every hour, the tags recorded the portion of time out of salt water (which we term "dry"). These dry-time percentages were used to determine haul-out status. Data were transmitted by satellite through the Argos system (CLS America, Inc., Largo, MD). There was limited storage on the SDRs. The instruments were not always able to upload data before storage capacity was filled, so there were missing data.

Our interest centered on factors that affect temporal haul-out patterns. The following explanatory variables were included in the model: (1) time-of-day as a categorical variable with 24 levels; (2) day-of-year as a continuous third-order polynomial variable ranging from 1 February to 31 July, and (3) interaction of time-of-day and day-of-year. Each seal was considered a random effect from the population of all seals.

A model for haul-out probability for each seal species was fitted based on data from seals with SDRs using a generalized linear mixed model that allowed for temporal autocorrelation using quasi-likelihood as described in [26]. Date and time-of-day, along with their interaction, had strong effects with P-values well below 0.05. Sex and age were initially considered as well, but were not significant with P-values above 0.05. Sex and age information was not available from aerial surveys, so in any case a pooled model over sex and age seemed reasonable. An example of the fitted haul-out model for ribbon seals from February through July is shown in Fig. 3.

The haul-out model is a generalized linear mixed model, just like the double-observer model in Section 2.3. The methods of [26] simply allow fast computing for very large datasets. All of the probability estimates and an estimated covariance matrix follow the same theory as in Section 2.3. The probability of haulout was fitted for all observed seals from Sections 2.2 and 2.3, denoted as $\hat{\mathbf{p}}_{h,k}$ for the kth species, with a corresponding estimated covariance matrix of $\hat{\mathbf{C}}_{h,k}$.

2.6. Combining detection and haul-out models

The detection models were developed for all species and all observers. Next, we combined detection and haul-out models, but for each species separately. The estimated probability of a seal being seen, both due to being hauled out of the water and detected from line transect sampling, for the kth species is

$$\hat{\mathbf{p}}_{dh,k} = \hat{\mathbf{p}}_{d,k} \odot \hat{\mathbf{p}}_{h,k}, \tag{5}$$

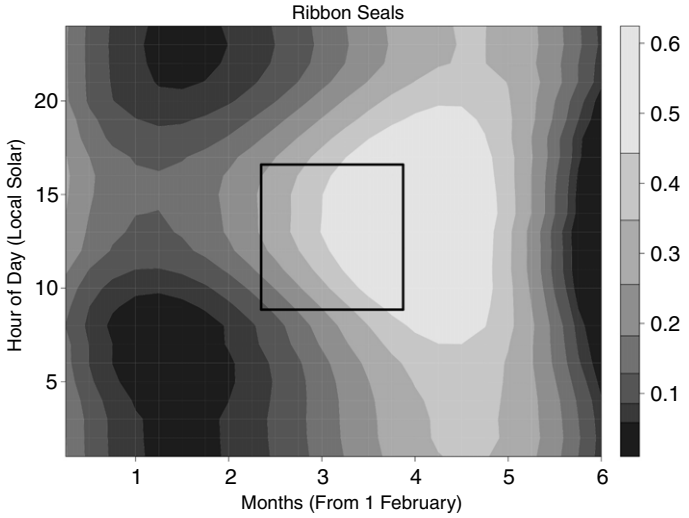

Fig. 3. A fitted model for the probability that a ribbon seal hauls out, as a function of date and time of day. The box with solid black lines shows the survey window, ranging from 13 April to 26 May, and from approximately 900 to 1600 local apparent time.

where $\hat{\mathbf{p}}_{d,k}$ is the subset of $\hat{\mathbf{p}}_d$ for the kth species, and the estimated variance covariance matrix of $\hat{\mathbf{p}}_{dh,k}$ is, from Appendix B,

$$\hat{\mathbf{C}}_{dh,k} = \hat{\mathbf{p}}_{d,k}\hat{\mathbf{p}}'_{d,k} \odot \hat{\mathbf{C}}_{h,k} + \hat{\mathbf{p}}_{h,k}\hat{\mathbf{p}}'_{h,k} \odot \hat{\mathbf{C}}_{d,k} - \hat{\mathbf{C}}_{h,k} \odot \hat{\mathbf{C}}_{d,k}$$

where $k = 1, 2, 3$ for the three species.

2.7. Estimating density in a spatial grid of plots

We created a grid of sample units for the analysis based on the AMSR ice data [13] (Polar Stereographic projection, cells = 12.5 km on a side). A sample unit for this analysis was defined as 4×4 AMSR grid cells (e.g., 16 total) forming sample units that were 50 km on a side. The final grid of sample units was made from $15 \times 12 = 192$ of these cells, numbered from 0 to 191 (Fig. 4). Rather than extrapolate the abundance estimate to the entire grid, we extrapolated only to the area within the minimum-convex polygon formed by the survey tracks (after subtracting land areas), as outlined by the heavy black lines in Fig. 4. For each sample unit above, average ice concentration was obtained by averaging the 16 cells of the AMSR ice data. These, in turn, were used to create a categorical variable of ice concentration for each grid cell in 25% class ranges (i.e., no ice, $0\%^{+}$–25% ice, $25\%^{+}$–50% ice, etc.).

Each sample unit was divided into three possible areas, as shown in Fig. 4. All land area, or area not contained in the minimum convex polygon, was eliminated, shown as the black area in Fig. 4. Any area falling within the sample unit that contained the observation strip of the line transect was computed for each survey and denoted $a_{t,i}$ for the ith sample unit during the tth date. Note that the distance bin 10° to 0 from horizontal was eliminated because sightings were very rare and species identification was unreliable. The total area of available habitat (non-land and all area inside the minimum convex polygon) is denoted A_i.

Let $D_{k,t,i}$ be the estimated density for species k in the ith sample unit during the tth date,

$$\hat{D}_{k,t,i} = \frac{1}{a_{t,i}} \sum_{j=1}^{n_{k,t,i}} \frac{s_{k,t,i,j}}{\hat{p}_{k,t,i,j}}, \tag{6}$$

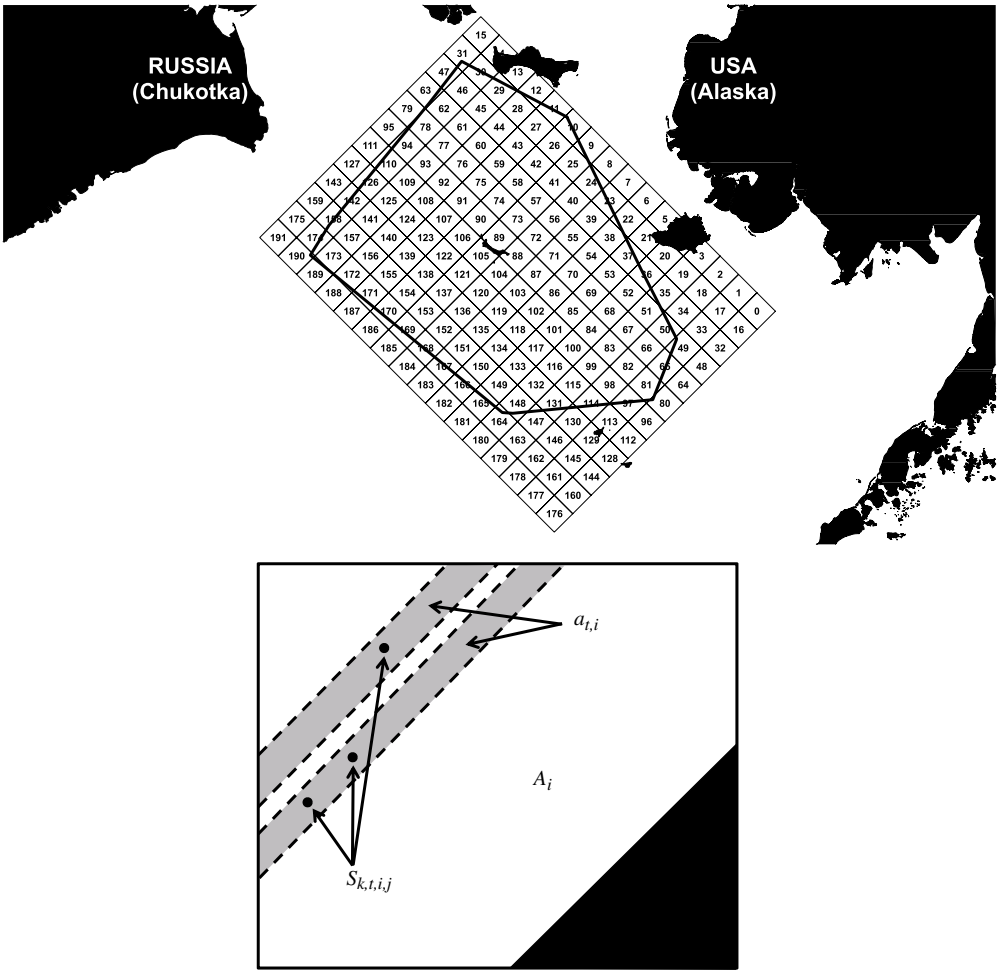

Fig. 4. Top: convex polygon of flights that includes all line transect surveys used in this analysis (all transects shown in Fig. 1, except the northern-most when ice conditions deteriorated). A grid of sample units was superimposed on this polygon to form sample units for estimating abundance within and spatial modeling among sample units. Bottom: example sampling unit, where $a_{t,i}$ is the area between the dashed lines from line-transect sampling that intersects the ith sampling unit for the tth date, and A_i is the total area of the ith sampling unit. Some area, shown in solid black, may be cut out due to land or the edge of the bounding polygon. Locations of observations are shown as solid dots.

where $s_{k,t,i,j}$ is the group size of the jth observation ($j = 1, 2, \ldots, n_{k,t,i}$) in the ith sample unit during the tth date, $p_{k,t,i,j}$ is the estimated haul-out/detection probability for the k, t, i, j group taken from (5), $a_{t,i}$ is the area of the line transect strip that fell within the ith sample unit during the tth date, and $n_{k,t,i}$ is the number of groups observed by line transect sampling in the ith sample unit during the tth date for the kth species. Eq. (6) is reminiscent of Horvitz–Thompson estimators, where observations are inflated by their probability of being sampled (observed in our case). It is convenient to write (6) as $\hat{\mathbf{d}}_{o,k} = \mathbf{U}_k\hat{\mathbf{s}}_k$ where $\hat{\mathbf{s}}_k$ is a vector with elements $s_{k,t,i,j}/(a_{t,i}\hat{p}_{k,t,i,j})$ for fixed k, and $\mathbf{U}_k$ contains zeros, except there are ones that indicate the sums in (6). The covariance matrix for all density estimates (6) for the kth species in spatial cells was obtained by noting that, using the delta method [15,25],

$$\hat{\text{var}}(\hat{\mathbf{s}}_k) \equiv \hat{\mathbf{C}}_{S,k} = \hat{\mathbf{S}}_k\hat{\mathbf{C}}_{dh,k}\hat{\mathbf{S}}_k,$$

where $\hat{\mathbf{S}}_k$ is a diagonal matrix with $\hat{s}_{k,t,i,j}/(a_{t,i}\hat{p}^2_{k,t,i,j})$ as the diagonal elements for fixed k, and then

$$\hat{\text{var}}(\hat{\mathbf{d}}_{o,k}) \equiv \hat{\mathbf{C}}_{o,k} = \mathbf{U}_k\hat{\mathbf{C}}_{s,k}\mathbf{U}_k'. \tag{7}$$

These data will now be used in a spatial hierarchical model.

3. Spatial hierarchical model

In the previous sections, we inflated individual point-referenced observations (6), accounting for the proportion of seals that were in the water and those that were out of the water but missed by observers. Next, these density estimates are used, along with their covariance matrix, in a spatial hierarchical model to make predictions throughout the grid for each time period. This model used relationships between seals and ice cover measured from satellite photography. In the terminology of [14], this is the process model that relies on ecological relationships.

Note that we expected the predictions to be highly imprecise, both individually and when summed over the grid for any given time period (because the samples are localized within the grid). However, one of the main concepts of this article is that these estimates can then be combined over time by assuming that the total abundance of a seal species is unchanged, just the spatial distribution changes among surveys. Part of the goal of the hierarchical model is an examination of estimates by time period to determine whether or not they can be combined into a single overall abundance estimate.

The data contained many zeros, so zero-inflated models were used. Let the vector of all seal densities in all cells be denoted $\mathbf{y}$, including any zeros. This process will be modeled as the product of a binary process $\mathbf{b}$ times an abundance process (expressed as density) $\mathbf{d}$, and hence $\mathbf{y} = \mathbf{b} \odot \mathbf{d}$. Each component was modeled separately, and we start by describing the density process. Each species will be modeled separately, so from now on we drop the subscript k.

3.1. Spatial density process

Let $\hat{\mathbf{d}}_o$ be the observed, non-zero densities as modeled through (6) and (7), where we drop the subscript k, and likewise $\hat{\text{var}}(\hat{\mathbf{d}}_o) = \hat{\mathbf{C}}_o$. This is a subset of all possible grid cells. Let us partition all grid cells into those that were observed as a subset of potentially observed density in all cells, $\hat{\mathbf{d}}_o \subseteq \mathbf{d}$. The data model us set up conditionally as a lognormal regression model;

$$\log(\hat{\mathbf{d}}_o)|\mathbf{m}_o \sim N(\mathbf{m}_o, \text{diag}(1/\hat{\mathbf{d}}_o)\hat{\mathbf{C}}_o\text{diag}(1/\hat{\mathbf{d}}_o)),$$

where $\text{diag}(1/\hat{\mathbf{d}}_o)\hat{\mathbf{C}}_o\text{diag}(1/\hat{\mathbf{d}}_o)$ comes from the delta method, and

$$M_{t,i} = \omega_0 + \sum_{j=1}^{J} \omega_j I(X_{t,i} \in S_j) + \delta_W W_{t,i}, \tag{8}$$

where $M_{t,i} \in \mathbf{m}_o$ (as in Section 1.1), $X_{t,i}$ is the percentage ice in the ith plot for the tth survey, S_j is the jth class interval for ice categories where, recall that $S_1 = (0, 25]$, $S_2 = (25, 50]$, $S_3 = (50, 75]$, $S_4 = (75, 100]$, $I(A)$ is the indicator function equal to one if A is true and 0 if A is false, $W_{t,i}$ is a random effect with a spatial conditionally-autoregressive (CAR) normal distribution [4] and [1, p. 79], and δ_W is a variance parameter. Note that (8) is applied to all grid cells, not just the observed ones, which allows prediction into unsampled grid cells. We assumed that the spatial CAR processes were independent across time periods, but shared common autocorrelation (ρ_W) and variance (δ_W) parameters. That is,

$$W_{i|t}|\mathbf{w}_{-i|t} \sim N(\xi_{W,i|t}, \zeta^2_{W,i|t}), \tag{9}$$

where $W_{i|t}$ is a random variable for the ith location for a given t and $\mathbf{w}_{-i|t}$ are conditional variables at all other locations for a given t, and

$$\xi_{W,i|t} = \frac{\rho_w}{|\mathcal{N}_i|} \sum_{j \in \mathcal{N}_{i|t}} W_{j|t},$$

and

$$\zeta^2_{W,i|t} = 1/|\mathcal{N}_i|,$$

with ρ_W the spatial autocorrelation parameter, $\mathcal{N}_i$ being the neighbors of i, and $|\mathcal{N}_i|$ the number of neighbors. We used the rook's move as neighbors; that is, if a site was not on an edge, it had four neighbors, one along each side of the grid cell.

3.2. Spatial binary process

Observed zeros did not mean that seals were absent from a spatial sample unit, because not all of the sample unit was surveyed. By changing the inflated observations to density, we have accounted for "extrapolation" to the whole sample unit for density given presence. This is not true for observed zeros. The use of survey area within sample units allowed us to model the probability of false negatives. Let the observed occurrence of any seal be denoted $B_{t,i}$ for the ith sample unit during the tth date, where $B_{t,i} = 1$ if any seal was observed (i.e., the density estimate was greater than 0), and 0 otherwise when that sample unit was surveyed but no seals were observed. In the space–time set of sample units, they could either have a density estimate, be zero, or be missing. A model is set up hierarchically,

$$[\mathbf{b}_o|\mathbf{q}_o, \boldsymbol{\chi}] = \prod_{t=1}^{27}\prod_{i=1}^{n_t}(Q_{t,i}\chi_{t,i})^{B_{t,i}}(1 - Q_{t,i}\chi_{t,i})^{1-B_{t,i}}, \tag{10}$$

where $\mathbf{b}_o$ is a vector of observed $\{B_{t,i}\}$, $\mathbf{q}_o$ is a vector of $\{Q_{t,i}\}$ for observed plots, which is a latent process of true presence/absence of a seal, and n_t is the number of plots that were sampled during the tth date. Note that if $Q_{t,i} = 0$, then $[B_{t,i}|Q_{t,i}] = [0|0] = 1$ is a degenerate distribution where $B_{t,i} = 0$ is the only outcome. However, if $Q_{t,i} = 1$, then the probability of observing a seal follows a Bernoulli distribution, where the probability of sighting is determined by the parameter $\chi_{t,i}$. Now this probability $\chi_{t,i}$ can be affected by the area surveyed on that plot through a generalized linear model,

$$\text{logit}(\chi_{i,j}) = \gamma_0 + \gamma_1 a_{t,i}, \tag{11}$$

recalling that $a_{t,i}$ is the area of the line transect strip that fell within the ith sample unit during the tth survey. The latent Bernoulli process $Q_{t,i}$ is the true presence/absence of any seal,

$$[\mathbf{q}|\mathbf{z}, \boldsymbol{\kappa}] = \prod_{t=1}^{27}\prod_{i=1}^{N_t}(\pi_{t,i})^{Q_{t,i}}(1 - \pi_{t,i})^{1-Q_{t,i}}, \tag{12}$$

where $\mathbf{q}$ is the vector of all $\{Q_{t,i}\}$ and N_t is the total number of plots that had any ice during the tth date. We now allow the latent Bernoulli process of true occurrence of seals to rely on ice cover and a spatial CAR model through a generalized linear mixed model with parameters $\boldsymbol{\kappa}$,

$$\text{logit}(\pi_{t,i}) = \kappa_0 + \sum_{j=1}^{J}\kappa_m I(X_{t,i} \in S_j) + \delta_Z Z_{t,i}, \tag{13}$$

where $X_{t,i}$, S_m and $I(A)$ are defined in (8), and $Z_{t,i}$ is random variable that follows a spatial conditional autoregressive (CAR) model as described in (9). Here, the spatial CAR model has variance parameter δ_Z and spatial autocorrelation parameter ρ_Z, and again allows prediction at unsampled grid cells.

3.3. Fitting the hierarchical model

All parameters in the hierarchical models described in Sections 3.1 and 3.2 were estimated using Markov Chain Monte Carlo (MCMC). The constructions given by Sections 3.1 and 3.2 allow us to obtain samples from the posterior distribution of,

$$[\mathbf{m}, \boldsymbol{\pi}, \boldsymbol{\omega}, \boldsymbol{\gamma}, \boldsymbol{\kappa}, \mathbf{z}, \mathbf{w}, \mathbf{c}, \mathbf{q}, \delta_W, \delta_Z, \rho_W, \rho_Z|\mathbf{d}_o, \mathbf{b}_o, \mathbf{X}, \mathbf{a}],$$

were $\mathbf{X} = \{X_{i,t}\}$, $\mathbf{a} = \{a_{i,t}\}$, $\mathbf{z} = \{Z_{i,t}\}$, $\mathbf{w} = \{W_{i,t}\}$, and $\mathbf{c} = \{\chi_{i,t}\}$. We used flat or very diffuse prior distributions for all remaining parameters. For the regression parameters on ice, we used $[\omega_j] = N(0, 10)$ in (8), $[\gamma_j] = N(0, 10)$ in (11), and $[\kappa_j] = N(0, 10)$ in (12) for all j. These priors were on the log and logit scales and were much broader than the posterior distributions but not so broad that they allowed computer overflows when exponentiating. Changing the priors had no influence on the results until they were much narrower, or so wide that they caused computational problems. The prior distributions on the spatial autocorrelation parameters ρ_W and ρ_Z were uniform on [0,1], and the spatial variance parameters δ_W and δ_Z were gamma(0.5, 2000) as suggested by Kelsall and Wakefield [18].

We used a burn-in of 10,000 samples, and then ran the MCMC chain until the coefficient of variation of the MCMC error, using the fixed-width analysis as suggested by Jones et al. [17], was less than 0.1% of the log of mean abundance for each species. That required 1,000,000 samples for bearded seals, 700,000 samples for ribbon seals, and 300,000 samples for spotted seals. Additionally, chains for all parameters were visually inspected, and we found no irregularities and all parameters had small MCMC errors relative to the standard deviation of their estimated values.

3.4. Abundance estimates

We take, as our estimate of abundance for the ith plot during the tth survey, the probability that seals are present in a plot, times the density given that seals are present in a plot, and then multiply this by the area of that plot. Hence, from (8) and (13),

$$\hat{Y}_{t,i} = \exp(M_{t,i})\pi_{t,i}A_i.$$

The estimates are totaled to provide an estimate of abundance for each survey,

$$\hat{Y}_{t,\bullet} = \sum_{i=1}^{N_A} \exp(M_{t,i})\pi_{t,i}A_i, \tag{14}$$

where $N_A = 192$ is the number of spatial grid cells.

A graph of each of these estimates is shown in Fig. 5A. The estimates (14) were computed at each iteration of MCMC, so the values shown in Fig. 5A were obtained as the median of the MCMC iterations, and the confidence intervals were obtained from the 2.5% and 97.5% quantiles of the MCMC iterations. For graphical presentation, the estimates and confidence interval endpoints were log-transformed. Note that we dropped the 27th survey. The diagnostics on MCMC convergence were poor for this date, and its variance was much higher than the rest.

After examining Fig. 5A, it seems that the abundance is relatively constant, and so it reasonable to combine all estimates through the survey period by averaging them to obtain an overall abundance estimate,

$$\hat{Y}_{\bullet,\bullet} = \frac{1}{N_T}\sum_{t=1}^{N_T}\sum_{i=1}^{N_A} \exp(M_{t,i})\pi_{t,i}A_i. \tag{15}$$

The posterior distributions of (15) for each species are shown in Fig. 5B, which were obtained from all MCMC estimates using a kernel density estimator in R with a bandwidth of 15,000. The mode of the marginal distribution was used as a point estimate with 2.5% and 97.5% quantiles of the MCMC sample for a 95% credible interval. The estimate for ribbon seals was 61,100 with a 95% credible interval of 35,200–189,300, the estimate for bearded seals was 61,800 with a 95% credible interval of 34,900–171,600, and the estimate for spotted seals was 233,700 with a 95% credible interval of 137,300–793,100.

3.5. Other parameter estimates

Our main goal was estimating abundance of all three seal species, which we achieved in the previous section. However, there are many other parameters of interest in the model, most of which

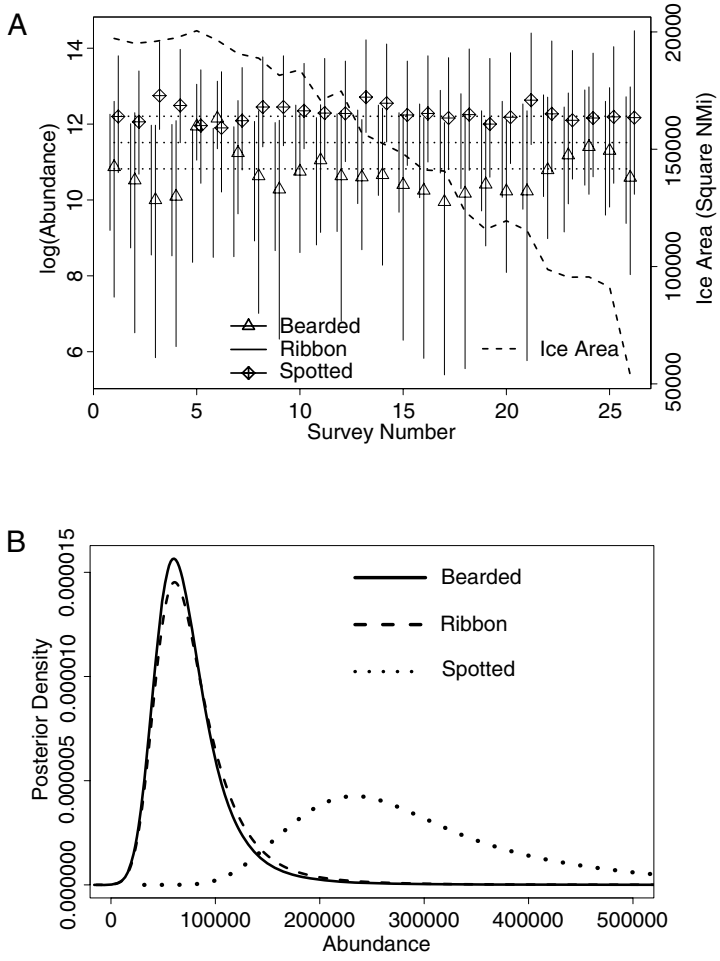

Fig. 5. A. Logarithm of the estimates of abundance for each survey date. The vertical bars are the 95% intervals. The estimates for each species were jittered on the x-axis for easier display. The total area of ice is also shown as a dashed line with the corresponding y-axis on the right. B. Posterior distributions of the abundance estimates averaged across survey dates for three species of ice seals.

will not be presented here, except for a brief summary. The two spatial autocorrelation parameters, for both density and binary processes, showed strong positive spatial autocorrelation with posterior distributions that had most of their mass near one. Likewise, the posterior distribution of γ_1 in (11) for all species was strongly positive with almost no values below zero in all MCMC samples, indicating, as one might expect, that increased area surveyed within a grid cell led to increased probability of detecting a seal.

Of greatest interest, besides the abundance estimates, was the relationship of each seal species to ice concentration. These relationships are shown in Fig. 6. There seems to be a partitioning of preferred ice cover habitat among the seals. The ribbon seals might be classified as having the most affinity to the ice edge, with their highest probability of occurrence in the 25%–50% ice cover class. Spotted seals have the highest probability of being in a sample unit overall, nearing 100% in the 25%–50% ice cover class, but its preferred ice cover class is 50%–75%. Bearded seals seem to be the most "interior" of the seal species with the highest probability of being in a sample in the 75%–100% ice cover class, with the highest densities there as well.

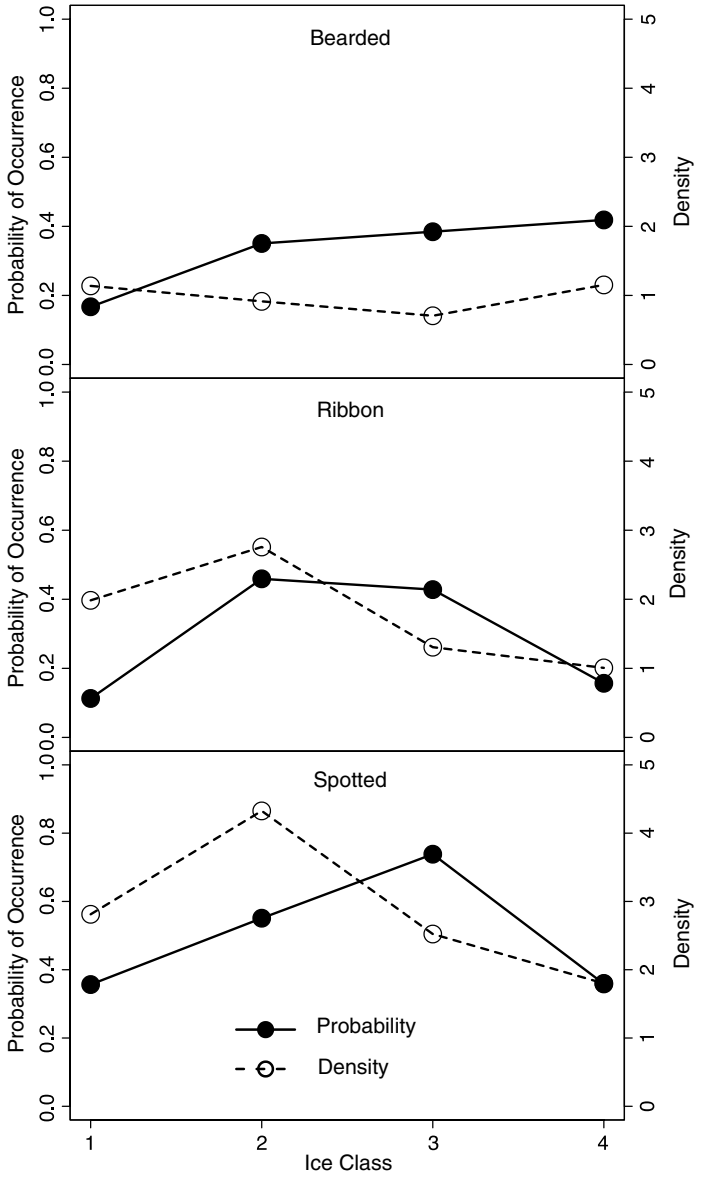

Fig. 6. Fitted estimates for the probability of at least one seal being in a sampling unit, and the density of seals given they are present, by ice class for each seal species.

3.6. Small simulation study

The hierarchical model, as described, is quite complicated. In order to test its efficacy (and our code), we simulated data. We selected plots 50:62, 66:78, 82:94, 98:110, 114:126, and 131:142 in Fig. 4A for survey dates 2:21, yielding 77 spatial plots for 20 time periods for 1540 total sample units. For each of the 20 time periods, we randomly selected one of all possible sets of 3×3 contiguous spatial plots; that is, we wanted to simulate spatially clustered sampling to mimic the real sampling. The 20 sets of 9 plots yielded 180 sampled units. For each sampled unit, we randomly

drew, with replacement, the area of the plot surveyed (the $a_{t,i}$ of Fig. 4B). The actual total plot area and observed ice coverage of each plot for each date was used for simulations. The presence of any seals in a plot was simulated using (13), by using the observed ice coverage values with $\boldsymbol{\kappa} = (0, -2.31, -0.185, -0.318, -1.787)$, which were the estimated values for the ribbon seals, and a random draw from the spatially autocorrelated $\{Z_{t,i}\}$, with $\delta_Z = 5.65$ and $\rho_Z = 0.947$, which were the estimated values for spotted seals. Using $\boldsymbol{\pi}$ back on the antilogit scale, the binary values were randomly drawn from $[\mathbf{q}|\mathbf{z}, \boldsymbol{\kappa}] = \text{Bern}(\boldsymbol{\pi})$. For the 180 sampled units, the probability of observing a seal, given at least one in the sample unit ($Q_{t,1} = 1$), was computed from (11) using $\alpha_0 = -7.06$ and $\alpha_1 = 3.51$, which were the estimated values from the bearded seals. Then, for all elements of $\mathbf{b}$ not in the sample $B_{t,i}$ was set to $Q_{t,i}$; for all $Q_{t,i} = 0$ in the sample $B_{t,i}$ was set to $Q_{t,i} = 0$; and for all $Q_{t,i} = 1$ in the sample, $B_{t,i}$ was randomly drawn from Bern(antilogit($\chi_{t,i}$)).

The density of seals in a plot was simulated using (8), by using the observed ice coverage values with $\boldsymbol{\omega} = (0, 0.856, 1.37, 0.854, 0.528)$, which were the estimated values for the spotted seals, and a random draw from the spatially autocorrelated $\{W_{t,i}\}$, with $\delta_W = 2.15$ and $\rho_W = 0.911$, which were again the estimated values for spotted seals. The lognormal values were created by taking $\mathbf{d} = \exp(\mathbf{m})$. Initially, the simulated abundance for each time period was $\sum_{i=1}^{77} D_{t,i} Q_{t,i} A_i$. After simulating several data sets, the average abundance was near 150,000, but with some variation among simulations and across the 20 time periods. Because one of our assumptions was a constant abundance across time periods, we proportionally adjusted each simulated $D_{t,i}$ for each time period t so that the total was exactly 150,000 for each time period. For the subset of plots that were sampled and had non-zero densities (say n_d of them), one draw was simulated from the first n_d rows and columns of the full multivariate normal distribution (covariance matrix taken from $\hat{\mathbf{C}}_{o,k}$ in (7)) that was estimated as the measurement error model for spotted seals. These measurement errors were added to the appropriate $\mathbf{m}_o \subset \mathbf{m}$ that were used as density data, $\mathbf{d}_o = \exp(\mathbf{m}_o)$, for the analysis. For the simulations, n_d ranged from 48 to 61 of the 180 sampled plots.

Four datasets were simulated and then abundance estimation used the same models and code as for the real data. The MCMC sampler used a burn-in of 40,000 samples (to allow for a bit slower convergence due to smaller sample sizes compared to the real data), and an additional 200,000 samples were used for estimation. No formal testing of MCMC was preformed because this was just a set of simulations, but all chains of all parameters were visually inspected and no irregularities were found. Log abundances were computed for each of the 20 time periods for each simulation, and are shown in Fig. 7A, which can be compared to Fig. 5A. The posterior densities of the four abundance estimates are shown in Fig. 7B. The thick black vertical line shows the true abundance of 150,000. This true abundance was within the 95% credibility intervals for all four posterior densities. Using the mean of the posterior density as an estimator, the four estimates were 218,000, 304,000, 229,000, 135,000, while using the median they were 197,400, 255,900, 195,000, 122,800, and using the mode they were 169,200, 210,200, 160,500, 110,200. While these simulations do not establish large-sample frequentist properties of the method (e.g., if credible intervals cover the true values controlling the simulation the correct proportion of time), it shows that the approach that we have taken seems reasonable.

4. Discussion and conclusions

Our approach used the data model, process model, and parameters framework described by Berliner [3] and outlined in Section 1.1 for this particular analysis. To summarize, we used maximum likelihood methods to estimate probabilities of observing seals, from both distance-sampling survey data (given they were on the ice) and satellite transmitter data (for the probability of being on the ice). The final data model also used an extension of Goodman's exact variance of a product [16] and the delta method [15] to take full account of modeling uncertainty. The process model then switched to a Bayesian spatial hierarchical model for combining density, detection and probability of presence for abundance estimation.

Fig. 5 provides the crucial argument for the approach we have taken in this article. In the Introduction, we described the problems of changing ice and a spatially and temporally dynamic population. The way we tackled this problem was to make many estimates over temporal scales

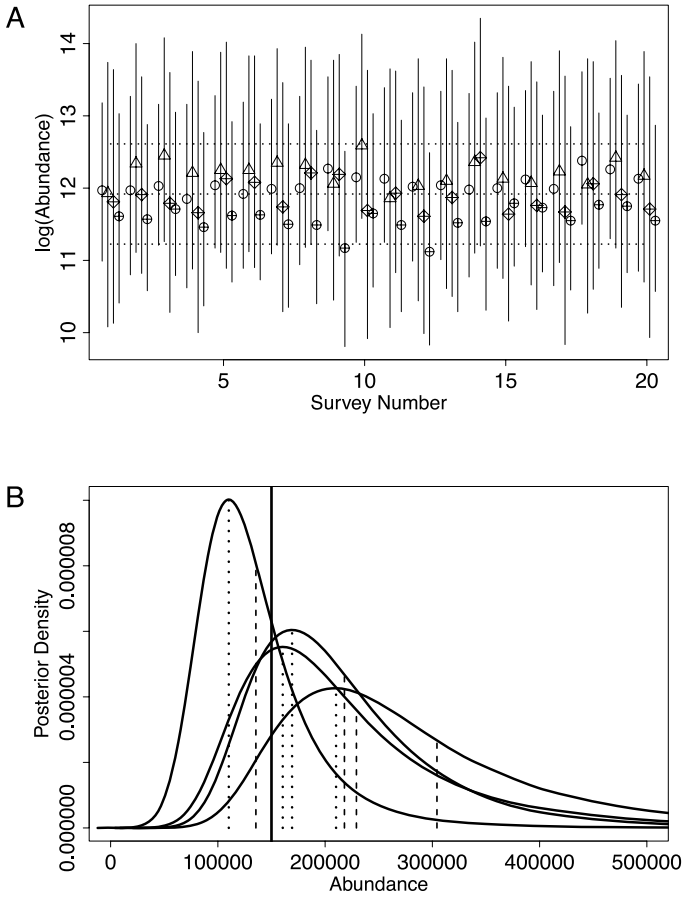

Fig. 7. A. Logarithm of the estimates of abundance for each survey time for four simulated datasets. The vertical bars are the 95% intervals. The estimates for each simulation were jittered on the x-axis for easier display. Horizontal dashed lines are at 75,000, 150,000 (true abundance), and 300,000. B. Posterior distributions of the abundance estimates averaged across times for four simulated data sets. Thick solid vertical line is true abundance, 150,000, for each simulation. Vertical dotted lines are posterior modes, and vertical dashed lines are posterior means.

that eliminate the problems of ice and seal movements. These estimates (14) are, admittedly, very imprecise. For example, the confidence intervals in Fig. 5A range from 400 to 1.2 million. However, Fig. 5A also reveals that, while ice was decreasing dramatically to a quarter of its original size, the abundance estimates remained relatively constant over time. This can be explained by three main factors: (1) the use of the haul-out model (Fig. 3), which can hold abundance constant even when the observed densities are changing, by adjusting observed densities due to the portion of the population that is in the water; (2) the use of regression on ice concentration, (8) and (13), which can remain relatively constant if a preferred ice type remains constant even when the overall ice changes, and (3) spatial autocorrelation, which mitigates some lack of fit from the regression model by propagating deviations from the fit throughout the grid cells for that time period. An example of the last effect occurs if, for the data from a specific date t, all of the observed density estimates were above the regression fit. In that case, spatial autocorrelation would cause predictions in nearby sample units to also be above their regression fits. Because the estimates remain relatively constant through time in Fig. 5A, we felt justified in combining them to get an estimate of abundance (15) that has better precision (Fig. 5B).

A small simulated dataset showed that inference for abundance, using the hierarchical model, was a reasonable approach. Note, however, we only simulated the "process" model, and not the data model. Simulating data for the data model would require simulating point patterns that we know little about. Methods for distance sampling have been extensively tested [9,10] and the haul-out model was essentially logistic regression with temporally autocorrelated random effects, which should not require testing. The overall data model was the product of probabilities that used an exact method, an extension of [16]. There were several uses of the delta method which is approximate, but sample sizes were fairly large so we assumed these approximations were quite accurate.

Finally, the results of this analysis are broadly consistent with what is known about the abundance and ecology of these species in the Bering Sea. For example, [11] estimated that there were about 90,000–100,000 ribbon seals in the Bering Sea. Our estimate of 61,100 ribbon seals in the central and eastern Bering Sea, an area that contains roughly half of the Bering Sea ribbon seal breeding habitat [8], is of correspondingly similar magnitude. Analogous comparisons for spotted and bearded seals also indicate that our estimates are of similar magnitude to previous estimates, though it must be emphasized the previous estimates for all three species were more typically in the realm of expert judgment than statistical estimation. The relationships between seal densities and classes of sea ice concentration (Fig. 6) mostly matched our expectations, especially that ribbon and spotted densities are highest in areas of moderate ice concentrations typical of the sea ice margins (e.g. [21]), whereas bearded seals are found in similar densities across all ice concentration classes, reflecting a more dispersed and somewhat more even distribution throughout the Bering Sea pack ice.

Acknowledgments

This project received financial support from NOAA's National Marine Fisheries Service, Alaska Fisheries Science Center. The authors thank Jeff Laake for sharing insights on distance sampling methods, and the point independence model in particular. Most of the bearded seal haul-out data were collected and made available by the Native Village of Kotzebue and the Alaska department of Fish and Game, with support from the Tribal Wildlife Grants Program of the US Fish and Wildlife Service (Grant Number U-4-IT). The satellite telemetry studies were conducted under the authority of Marine Mammal Protection Act Scientific Research Permits 15126, 782-1765, 782-1676, and 358-1787. The National Science Foundation provided crucial in kind support for vessel and helicopter operations used in the survey. The authors wish to thank an anonymous guest editor and reviewer for helpful suggestions.

Appendix A. Modeling the detection function

To fit the model (3) we use maximum likelihood by minimizing the negative log-likelihood,

$$\mathcal{L}(\boldsymbol{\theta}|\{\mathcal{D}_{k,j}\}) = -\log\left(\prod_{k,j} p(\mathcal{D}_{k,j}|\boldsymbol{\theta})\right). \tag{A.1}$$

Eq. (A.1) was minimized for $\log(\sigma)$ and $\log(\lambda)$ to ensure they remained positive. The fitted models are shown in Fig. 2. Confidence intervals were estimated by using the Hessian, $\mathbf{H}$, at the minimized value of (A.1), and assuming $\hat{\boldsymbol{\theta}} \sim N(\boldsymbol{\theta}, \mathbf{H}^{-1})$. The estimated value for $\hat{\lambda}$ (after back-transforming) was 0.884 with a 95% confidence interval of 0.772–1.013 (by back-transforming the endpoints of the confidence interval), indicating it was much different than a half-normal, which has $\lambda = 0.5$, justifying the use of the exponential power series model. Detection for bearded seals ($\tau_1 \equiv 0$) was highest, with lower values for ribbon seals ($\hat{\tau}_2 = -0.616$, 95% confidence interval -0.795–-0.438) and spotted seals ($\hat{\tau}_3 = -0.441$, 95% confidence interval -0.586–-0.296). Detection increased with larger groups of seals ($\hat{\beta}_1 = 0.060$, 95% confidence interval -0.013–0.132), and although the confidence interval included zero, we retained group size in our model because we believe it is important.

From the estimated parameters we need the fitted detections for all observations, denoted as the vector $\hat{\mathbf{p}}_v$, and its covariance matrix. Note the when fitting linear models with $\hat{\boldsymbol{\theta}}$ estimated regression

parameters, a fit such as $\mathbf{X}\hat{\boldsymbol{\theta}}$ has a variance of $\mathbf{X}\mathbf{C}_\theta\mathbf{X}'$ where $\mathbf{X}$ is the design matrix and $\mathbf{C}_\theta$ is the covariance matrix of $\hat{\boldsymbol{\theta}}$; e.g. $\mathbf{C}_\theta = (\mathbf{X}\boldsymbol{\Sigma}^{-1}\mathbf{X}')^{-1}$. Here, the fits are nonlinear functions of the parameters. Let $l_{\mathcal{D}|k,j}$ be the lower bin cutoff and $u_{\mathcal{D}|k,j}$ be the upper bin cutoff for the k, jth observation. Then the fitted detection for the k, jth observation was the average value of $g(x)$ in its bin class,

$$\hat{p}_{k,j} = \frac{\int_{l_{\mathcal{D}|k,j}}^{u_{\mathcal{D}|k,j}} g_{k,j}(x|\hat{\boldsymbol{\theta}})dx}{u_{\mathcal{D}|k,j} - l_{\mathcal{D}|k,j}}, \tag{A.2}$$

where the integrals were obtained numerically. In order to obtain the covariance matrix of all of the fits jointly we will use a multivariate version of delta method [15,25]. In general, let $(f_1(\mathbf{x}), f_2(\mathbf{x}), \ldots, f_n(\mathbf{x}))'$ be a vector of functions, where each $f_i(\mathbf{x})$ is a function of a vector of random variables $\mathbf{x}$ with $E[\mathbf{x} = \boldsymbol{\mu}]$ and $\mathrm{cov}(\mathbf{x}) = \mathbf{C}_x$. Then using a Taylor Series expansion,

$$\begin{pmatrix} f_1(\mathbf{x}) \\ f_2(\mathbf{x}) \\ \vdots \\ f_n(\mathbf{x}) \end{pmatrix} = \begin{pmatrix} f_1(\boldsymbol{\mu}) \\ f_2(\boldsymbol{\mu}) \\ \vdots \\ f_n(\boldsymbol{\mu}) \end{pmatrix} + \begin{pmatrix} \frac{\partial f_1(\mathbf{x})}{\partial x_1} & \frac{\partial f_1(\mathbf{x})}{\partial x_2} & \cdots & \frac{\partial f_1(\mathbf{x})}{\partial x_p} \\ \frac{\partial f_2(\mathbf{x})}{\partial x_1} & \frac{\partial f_2(\mathbf{x})}{\partial x_2} & \cdots & \frac{\partial f_2(\mathbf{x})}{\partial x_p} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_n(\mathbf{x})}{\partial x_1} & \frac{\partial f_n(\mathbf{x})}{\partial x_2} & \cdots & \frac{\partial f_n(\mathbf{x})}{\partial x_p} \end{pmatrix} (\mathbf{x} - \boldsymbol{\mu}) + \cdots,$$

where higher order terms are ignored. For short-hand, we write this as $\mathbf{f}(\mathbf{x}) \approx \mathbf{f}(\boldsymbol{\mu}) + \boldsymbol{\Delta}(\mathbf{x} - \boldsymbol{\mu})$. Following the usual development of the delta method [15,25], $\mathbf{f}(\mathbf{x}) - \mathbf{f}(\boldsymbol{\mu}) \approx \boldsymbol{\Delta}(\mathbf{x} - \boldsymbol{\mu})$ and so $E[\mathbf{f}(\mathbf{x}) - \mathbf{f}(\boldsymbol{\mu})][\mathbf{f}(\mathbf{x}) - \mathbf{f}(\boldsymbol{\mu})]' \approx \boldsymbol{\Delta}\mathbf{C}_x\boldsymbol{\Delta}'$. In our application, $\hat{\boldsymbol{\theta}}$ takes the place of $\mathbf{x}$ with covariance matrix $\mathbf{H}^{-1}$, so we only need partial derivatives of the average detection function for a bin class,

$$\frac{\partial}{\partial\theta_i}\frac{\int_l^u g(x|\boldsymbol{\theta})}{u - l} = \frac{\partial}{\partial\theta_i}\frac{\int_l^u \exp\left(-\left(\frac{x}{\exp(\sigma_{\log,k,j})}\right)^{\exp(-\lambda_{\log})}\right)dx}{u - l},$$

where $\lambda_{\log}$ is the natural log of λ and $\sigma_{\log,k,j} = \beta_0 + \tau_k + \beta_1 s_{k,j}$ was modeled as in (2). We obtain

$$\frac{\partial}{\partial\boldsymbol{\theta}}\frac{\int_{l_{\mathcal{D}|k,j}}^{u_{\mathcal{D}|k,j}} g_{k,j}(x|\boldsymbol{\theta})dx}{u_{\mathcal{D}|k,j} - l_{\mathcal{D}|k,j}} = \begin{pmatrix} v_1(\beta_0, \beta_1, \tau_k, \lambda_{\log}, u_{\mathcal{D}|k,j}, l_{\mathcal{D}|k,j}), \\ s_{k,j}v_1(\beta_0, \beta_1, \tau_k, \lambda_{\log}, u_{\mathcal{D}|k,j}, l_{\mathcal{D}|k,j}), \\ I(k=2)v_1(\beta_0, \beta_1, \tau_k, \lambda_{\log}, u_{\mathcal{D}|k,j}, l_{\mathcal{D}|k,j}), \\ I(k=3)v_1(\beta_0, \beta_1, \tau_k, \lambda_{\log}, u_{\mathcal{D}|k,j}, l_{\mathcal{D}|k,j}), \\ v_2(\beta_0, \tau_k, \beta_1, \lambda_{\log}, u_{\mathcal{D}|k,j}, l_{\mathcal{D}|k,j}), \end{pmatrix} \equiv \mathbf{v}_{k,j},$$

for $\boldsymbol{\theta} = (\beta_0, \beta_1, \tau_2, \tau_3, \lambda_{\log})'$, where

$$v_1(\beta_0, \tau_k, \beta_1, \lambda_{\log}, u, l) = \frac{\int_l^u g(x|\boldsymbol{\theta})\left(\frac{x}{\exp(\sigma_{\log,k,j})}\right)^{\exp(-\lambda_{\log})}dx}{\exp(\lambda_{\log})(u - l)}$$

and

$$v_2(\beta_0, \tau_k, \beta_1, \lambda_{\log}, u, l) = \frac{\int_l^u g(x|\boldsymbol{\theta})\left(\frac{x}{\exp(\sigma_{\log,k,j})}\right)^{\exp(-\lambda_{\log})}\log\left(\frac{x}{\exp(\sigma_{\log,k,j})}\right)dx}{\exp(\lambda_{\log})(u - l)}.$$

Note that $\exp(\lambda_{\log}) = \lambda$, but the Hessian $\mathbf{H}$, from the minimization of (A.1), is in terms of $\lambda_{\log}$. Then, using the multivariate version of the delta method [15,25], the covariance matrix of all fitted values in (A.2) is approximated by

$$\mathrm{var}(\hat{\mathbf{p}}_v) = \begin{pmatrix} \mathbf{v}'_{1,1} \\ \mathbf{v}'_{1,2} \\ \vdots \\ \mathbf{v}'_{3,n_3} \end{pmatrix} \mathbf{H}^{-1} \begin{pmatrix} \mathbf{v}_{1,1}, & \mathbf{v}_{1,2}, & \ldots, & \mathbf{v}_{3,n_3} \end{pmatrix} \equiv \mathbf{C}_v.$$

Appendix B. Covariance of the product of random variables

Goodman [16] shows a useful result on the exact variance of the product of two random variables. Let X and Y be two random variables with $E[X] = \mu_x$, $E[Y] = \mu_y$ and X and Y independent of each other, then

$$\text{var}(XY) = \mu_x^2\text{var}(Y) + \mu_y^2\text{var}(X) + \text{var}(X)\text{var}(Y). \tag{B.1}$$

Bohrnstedt and Goldberger [5] generalized the result of [16] to covariances of two products. Let X_1, X_2, Y_1 and Y_2 be random variables with $E[X_1] = \mu_{x_1}$, $E[Y_1] = \mu_{y_1}$, $E[X_2] = \mu_{x_2}$ and $E[Y_1] = \mu_{y_1}$ with X_i independent of Y_j for all i and j, but X_1 and X_2 may be correlated as well as Y_1 correlated to Y_2. Then, following [5] we obtain the following,

$$\begin{aligned}\text{cov}(X_1Y_1, X_2Y_2) &= E[(X_1Y_1 - \mu_{x_1}\mu_{y_1})(X_2Y_2 - \mu_{x_2}\mu_{y_2})] \\ &= \mu_{x_1}\mu_{x_2}\text{cov}(Y_1, Y_2) + \mu_{y_1}\mu_{y_2}\text{cov}(X_1, X_2) + \text{cov}(X_1, X_2)\text{cov}(Y_1, Y_2). \end{aligned}\tag{B.2}$$

Note that if $X_1 = X_2$ and $Y_1 = Y_2$, then (B.2) equals (B.1), as it should. A direct product of two vectors will contain many product variances and covariances, which will be a matrix with elements given by results (B.1) and (B.2). More generally, let $\mathbf{x}$ be a vector of random variables with $E(\mathbf{x}) = \boldsymbol{\mu}_x$ and $\text{var}(\mathbf{x}) = \mathbf{V}_x$ and let $\mathbf{y}$ be a vector of random variables with $E(\mathbf{y}) = \boldsymbol{\mu}_y$ and $\text{var}(\mathbf{y}) = \mathbf{V}_y$, and let $\mathbf{x}$ be independent of $\mathbf{y}$. Then,

$$\text{var}(\mathbf{x} \odot \mathbf{y}) = (\boldsymbol{\mu}_x\boldsymbol{\mu}_x') \odot \mathbf{V}_y + (\boldsymbol{\mu}_y\boldsymbol{\mu}_y') \odot \mathbf{V}_x + \mathbf{V}_x \odot \mathbf{V}_y, \tag{B.3}$$

where $\odot$ is the direct (Hadamard, element-wise) product.

B.1. Unbiased estimation

Goodman [16] also shows that if $\hat{\text{var}}(X)$ is unbiased for $\text{var}(X)$ and $\hat{\text{var}}(Y)$ is unbiased for $\text{var}(Y)$, then an unbiased estimate of (B.1) is

$$\begin{aligned}\hat{\text{var}}(XY) &= [X^2 - \hat{\text{var}}(X)]\hat{\text{var}}(Y) + [Y^2 - \hat{\text{var}}(Y)]\hat{\text{var}}(X) + \hat{\text{var}}(X)\hat{\text{var}}(Y) \\ &= X^2\hat{\text{var}}(Y) + Y^2\hat{\text{var}}(X) - \hat{\text{var}}(X)\hat{\text{var}}(Y), \end{aligned}\tag{B.4}$$

because $E(X^2) = \text{var}(X) + \mu_x^2$ by the definition of variance. For the covariance case, if $\hat{\text{cov}}(X_1, X_2)$ is unbiased for $\text{cov}(X_1, X_2)$ and $\hat{\text{cov}}(Y_1, Y_2)$ is unbiased for $\text{cov}(Y_1, Y_2)$, then an unbiased estimate of (B.2) is

$$\begin{aligned}\hat{\text{cov}}(X_1Y_1, X_2, Y_2) &= [X_1X_2 - \hat{\text{cov}}(X_1, X_2)]\hat{\text{cov}}(Y_1, Y_2) \\ &\quad + [Y_1Y_2 - \hat{\text{cov}}(Y_1, Y_2)]\hat{\text{cov}}(X_1, X_2) + \hat{\text{cov}}(X_1, X_2)\hat{\text{cov}}(Y_1, Y_2) \\ &= X_1X_2\hat{\text{cov}}(Y_1, Y_2) + Y_1Y_2\hat{\text{cov}}(X_1, X_2) - \hat{\text{cov}}(X_1, X_2)\hat{\text{cov}}(Y_1, Y_2), \end{aligned}\tag{B.5}$$

because $E(X_1X_2) = \text{cov}(X_1, X_2) + \mu_{x_1}\mu_{x_2}$ by the definition of covariance. Generalizing to the vector case, suppose $\hat{\mathbf{V}}_x$ is unbiased for $\text{var}(\mathbf{x})$ and $\hat{\mathbf{V}}_y$ is unbiased for $\text{var}(\mathbf{y})$, then an unbiased estimate of (B.3) is

$$\hat{\text{var}}(\mathbf{x} \odot \mathbf{y}) = (\mathbf{x}\mathbf{x}') \odot \hat{\mathbf{V}}_y + (\mathbf{y}\mathbf{y}') \odot \hat{\mathbf{V}}_x - \hat{\mathbf{V}}_x \odot \hat{\mathbf{V}}_y. \tag{B.6}$$

References

[1] S. Banerjee, B.P. Carlin, A.E. Gelfand, Hierarchical Modeling and Analysis for Spatial Data, Chapman & Hall/CRC, New York, New York, USA, 2004.
[2] D. Bates, M. Maechler, Lme4: Linear mixed-effects models using S4 classes, r package version 0.999375-32, 2009.
[3] L.M. Berliner, Hierarchical Bayesian time series models, in: K. Hanson, R. Silver (Eds.), Maximum Entropy and Bayesian Methods, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1996, pp. 15–22.
[4] J. Besag, J. York, A. Mollié, Bayesian image restoration, with two applications in spatial statistics, Annals of the Institute of Statistical Mathematics 43 (1991) 1–20.

[5] G.W. Bohrnstedt, A.S. Goldberger, On the exact covariance of products of random variables, Journal of the American Statistical Association 64 (1969) 1439–1442.
[6] D.L. Borchers, J.L. Laake, C. Southwell, C.G.M. Paxton, Accommodating unmodeled heterogeneity in double-observer distance sampling surveys, Biometrics 62 (2006) 372–378.
[7] D.L. Borchers, W. Zucchini, R.M. Fewster, Mark-recapture models for line transect surveys, Biometrics 54 (1998) 1207–1220.
[8] P.L. Boveng, J.L. Bengtson, T. Buckly, M.F. Cameron, S.P. Dahle, B.A. Megrey, J.E. Overland, N.J. Williamson, Status review of the ribbon seal (*Histriophoca fasciata*), Tech. Rep., U.S. Dept. of Commerce, NOAA Technical Memorandum NMFS-AFSC-191, 2008.
[9] S.T. Buckland, D.R. Anderson, K.P. Burnham, J.L. Laake, D.L. Borchers, L. Thomas, Introduction to Distance Sampling, Oxford University Press, Oxford, 2001.
[10] S.T. Buckland, D.R. Anderson, K.P. Burnham, J.L. Laake, D.L. Borchers, L. Thomas, Advanced Distance Sampling: Estimating Abundance of Biological Populations, Oxford University Press, Oxford, 2004.
[11] J.J. Burns, Ribon seal phoca fasciata Zimmerman, 1783, in: S.H. Ridgway, R. Harrison (Eds.), Handbook of Marine Mammals, Vol. 2: Seals, Academic Press, London, 1981, pp. 89–109.
[12] M.F. Cameron, P.L. Boveng, Abundance and Distribution surveys for ice seals aboard USCG Healy and the Oscar Dyson, AFSC Quarterly Research Reports April–June, 2007, pp. 12–14.
[13] D. Cavalieri, T. Markus, J. Comiso, AMSR-E/Aqua Daily L3 12.5 km Brightness Temperature, Sea Ice Concentration, and Snow Depth Polar Grids V002, March–June, 2007, National Snow and Ice Data Center, Digital Media, Boulder, Colorado, USA, 2004, updated daily.
[14] N. Cressie, K.A. Calder, J.S. Clark, J.M. Ver Hoef, C.K. Wikle, Accounting for uncertainty in ecological analysis: the strengths and limitations of hierarchical statistical modeling, Ecological Applications 19 (2009) 553–570.
[15] R. Dorfman, A note on the delta-method for finding variance formulae, The Biometric Bulletin 1 (1938) 129–137.
[16] L.A. Goodman, On the exact variance of products (Corr: V57 P917), Journal of the American Statistical Association 55 (1960) 708–713.
[17] G.L. Jones, M. Haran, B.S. Caffo, R. Neath, Fixed-width output analysis for Markov Chain Monte Carlo, Journal of the American Statistical Association 101 (2006) 1537–1547.
[18] J. Kelsall, J. Wakefield, Discussion of: "Bayesian models for spatially correlated disease and exposure data" by N.G. Best, L.A. Waller, A. Thomas, E.M. Conlon, R. Arnold, in: J. Bernardo, J. Berger, A. Dawid, A. Smith (Eds.), in: Bayesian Statistics, vol. 6, Oxford University Press, 1999, p. 151.
[19] J. Laake, D. Borchers, Methods for incomplete detection at distance zero, in: S. Buckland, D. Anderson, K. Burnham, J. Laake, L. Thomas (Eds.), Advanced Distance Sampling, Oxford University Press, 2004, pp. 108–189.
[20] R.C. Littell, R.C. Milliken, W.W. Stroup, R. Wolfinger, SAS System for Mixed Models, SAS Publishing, Cary, North Carolina, 1996.
[21] L.F. Lowry, V.N. Burkanov, K.J. Frost, M.A. Simpkins, R. David, D.P. DeMaster, R. Suydam, A. Springer, Habitat use and habitat selection by spotted seals (*Phoca largha*) in the Bering sea, Canadian Journal of Zoology 78 (2000) 1959–1971.
[22] K.H. Pollock, A family of density estimators for line-transect sampling, Biometrics 32 (1978) 475–478.
[23] P.X. Quang, E.F. Becker, Combining line transect and double count sampling techniques for aerial surveys, Journal of Agricultural, Biological, and Environmental Statistics 2 (1997) 230–242.
[24] R Development Core Team R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, ISBN: 3-900051-07-0, 2010.
[25] J.M. Ver Hoef, Who invented the delta method?, The American Statistician 66 (2012) 124–127.
[26] J.M. Ver Hoef, J.M. London, P.L. Boveng, Fast computing of some generalized linear mixed pseudo-models with temporal autocorrelation, Computational Statistics 25 (2010) 39–55. http://dx.doi.org/10.1007/s00180–009–0160–1.