Spatio-temporal modeling of relative animal density using a long time series of line transect surveys with clustering and censoring

G. H. Givens*,†

Abstract

We introduce a probability model for data arising from aerial line transect surveys, with the goal of estimating relative animal density. This model includes consideration of animal clustering and censored observations due to effort truncation and flights with zero animal sightings. It also includes terms for spatio-temporal covariates that affect detection probabilities and animal presence. Approximate model fitting is accomplished using generalized additive modeling techniques for censored data. Estimation of uncertainty relies on bootstrapping. This model was motivated by study of a large, long-term dataset of bowhead whale surveys in the western Arctic. The fitted model is shown to map the spatio-temporal pattern of the fall bowhead migration effectively. Additional model terms may be added to test for potential localized zones of unusual scarcity or abundance such as what could be introduced by avoidance of industrial activities or hunting, or by variation in prey availability, whale behavior, or other environmental factors. Although localized scarcity phenomena are supported here by significant statistical evidence, the causes are not well understood.

*Department of Statistics, 1877 Campus Delivery, Colorado State University, Fort Collins, CO 80523 U.S.A. geof@lamar.colostate.edu

†This work was supported by the North Slope Borough (Alaska) and the National Oceanic and Atmospheric Administration (through the Alaska Eskimo Whaling Commission). Craig George, Robert Suydam, Chuck Monnett, Jennifer Hoeting, Laura Beri, Janet Clarke, Dave Rugh, and Robert Flynn for helpful discussions and advice. Megan Ferguson is greatly thanked for participating in a thoughtful and intellectually challenging dialog that helped to refine many aspects of this paper.

1 Introduction

In many wildlife management and conservation biology problems, it is important to estimate an index of relative animal abundance. This estimate, indexed over space and/or time, describes the number of animals in any spatio-temporal region up to a proportionality constant. Relative abundance indices are important for understanding the spatio-temporal distribution of animals and can capture such signals as local abundance due to abundant feeding resources, aggregations in a breeding region, and migratory patterns.

As suggested by the examples of feeding, breeding and migration, relative abundance may be strongly related to spatial, temporal or other covariates. Key to understanding the ecological and management issues for the surveyed population is the ability to link the relative abundance index to such external covariates.

We focus here on abundance data obtained from single-observer aerial line-transect surveys. Alone, such data can be used to estimate the marginal relative abundance. With records of covariate data (usually contemporaneous and localized), the abundance index can be related to the effects of such variables. Furthermore, the nature of the relationships between covariates and abundance is often itself a primary scientific question. For example, consider an aerial survey of animals over a region of variable habitat. Animal presence may be strongly affected by food resources that vary with vegetation across the survey area. The ease of detecting animals may also vary, due to the variation in vegetative cover. To fully understand relative abundance then, estimates should adjust for the effect of vegetation on sighting and also facilitate understanding of how animal presence depends on habitat.

This paper is divided in three main parts. First, we develop a likelihood function for situations where data are obtained from a collection of line transect surveys and spatio-temporal covariates. Although this likelihood is comprehensive and could be numerically maximized in principle, it is sufficiently complex that more pragmatic estimation strategies are attractive. Second, we illustrate a simplification of the general model that can be used to draw inference about spatio-temporal relative abundance and covariate effects. Finally, we apply the method to a large dataset for bowhead whales. By no means is this application simple: it provides a realistic example that has important implications for the scientific understanding, conservation and management of these whales.

1.1 Line transect surveys

The models here relate to line transect surveys. To conduct a line transect survey, a survey region is selected. Interest in relative abundance is limited to this region, which is sometimes believed to cover the entire range of a distinct subpopulation, although this latter assumption is not necessary. The survey effort is conducted along a sequence of line segments superimposed over the region. These 'tracklines' must be laid down randomly over the region with respect to the distribution of animals. The surveyor then travels along the tracklines, trying to detect animals. Each sighting is recorded, along with the animal's perpendicular distance from the trackline, the observer's location along the trackline, and any covariates.

Since detection of animals generally decreases with their distance from the trackline, the relative abundance of sightings along the trackline must be corrected for sighting distances. Adjustments rely on a 'detection function', g, for which $g(y)$ denotes the probability of sighting an animal given that it is located at distance y from the trackline. Except for special types of surveys, one assumes $g(0) = 1$. Let w denote the distance from the trackline beyond which it is reasonable to assume that detection is effectively precluded, or the distance beyond which the data are truncated by removing more distant observations from analysis. Truncation of distant cases is often reasonable because estimation of the shape of g can be sensitive to such outliers.

Suppose that one wishes to produce relative abundance estimates for two portions of the survey region, with areas A_1 and A_2, with corresponding total on-effort transect lengths of L_1 and L_2. Then, the corresponding areas where sightings could have occurred are given by $a_i = 2wL_i$ for $i = 1, 2$. If all n_i animals are detected in A_i then under certain assumptions including issues of randomness and measurement error (see Buckland et al., 2001) the density of animals within A_i is the number of animals per unit area, $n_i/2wL_i$.

Of course, only a portion of all available animals are sighted. In this case, n_i denotes the number of sightings. Let P_i denote the proportion of animals in region A_i that are sighted. Then adjustment for the proportions of sighted animals produces density estimates of $n_i/2wL_iP_i$. Finally, making a further adjustment for the sizes of A_1 and A_2, we obtain the relative abundances of $D_i = A_in_i/2wL_iP_i$.

In practice, counts like n_i are observations of random variables, and P_i must be estimated from the sighting data. The unconditional probability of detecting an animal in a strip of width $2w$ and length L_i is $P_i = \frac{1}{w}\int_0^w g(y)\,dy$. Estimation of P_i boils down to estimation of g. Thomas et al. (2006) and Buckland et al. (2001; 2004) describe a method for estimation of g based on the collection of observed sighting distances.

1.2 Whale data

The purpose of this paper is to develop and apply appropriate spatial line transect methods for analysis of a dataset pertaining to migratory whales in the western Arctic, namely the population of bowhead whales (*Balaena mysticetus*) in the Bering, Chukchi, and Beaufort Seas. A key goal is to produce a spatio-temporal map of relative abundance.

The bowhead whale is a migratory species. The population studied here winters in the Bering Sea and generally follows the Chukotkan and Alaskan coasts northeast during the spring. In spring and early summer the whales pass through the Chukchi Sea. Most continue east to summer in the eastern Beaufort Sea feeding grounds. As fall begins, the bowheads return west and generally cross the Chukchi Sea to reach the north coast of the Chukotkan Peninsula before traveling along that coast back to their wintering grounds. Descriptions of the migration are given by Rugh et al. (2003) and Moore and Reeves (1993).

In 1979 the Bowhead Whale Aerial Survey Project (BWASP) was initiated by the US Bureau of Land Management to study the possible effects of industrial activity on the marine and coastal environment in the Beaufort Sea. A proposed oil and gas lease sale prompted

this project, and industry impacts continue to be a motivation in recent years. Responsibility for the survey has shifted to the US Minerals Management Service (MMS), the Naval Ocean Systems Center and affiliated MMS contractors.

The BWASP dataset includes records of many aerial survey flights each year from 1979 to 2008, ranging through the Chukchi and Beaufort Seas and during all portions of spring, summer, and fall. Most flights occurred in fall over the Beaufort Sea.

The BWASP data can be used to address several important concerns about bowhead whales and resource management. Research about these whales is hampered by challenging weather and ice conditions, with only a limited portion of the year providing both whale presence and suitable observation conditions. Although many migratory observations and some tracking data exist, a comprehensive and precise spatio-temporal map has been beyond reach. Thus, the first goal of our BWASP analysis is to map the migration. Beyond the descriptive value of such a map, it will also useful as a baseline for future study: as the Arctic continues warm and sea ice retreats, the impacts of changing environmental conditions and human activity must be compared to earlier conditions.

Second, BWASP data are valuable to address issues related to bowhead hunting by native Inupiat and Yupik communities in northern and western Alaska, who rely on limited hunting quotas to satisfy their subsistence and cultural needs. This hunting is regulated by the International Whaling Commission. Changes to bowhead migratory patterns could reduce availability of whales to the hunters and require villagers to venture greater distances at greater risk. It is of concern, therefore, to assess whether any deviations from the standard migratory path might be correlated with anthropogenic impacts or other factors. Impacts could include spatio-temporal displacement or behavioral changes (e.g., swimming unusually quickly through a region). Aside from acute phenomena, areas of low apparent whale density may be attributable to a variety of other factors like prey availability or other environmental factors.

Previous use of the BWASP data has been limited. Manly et al. (2007) explored how human activities affected bowhead distribution using a Poisson regression approach applied to only 1996-1998 data. They found evidence that seismic activity is associated with reduced abundance and perhaps with swimming direction. Schick and Urban (2000) showed that bowhead whale distribution patterns were correlated with distance to oil rigs. Treacy et al. (2006) sought to link annual variation in ice distribution to migratory patterns.

A detailed description of the BWASP survey methods is given by Treacy (2002) and Monnett and Treacy (2005). Our brief summary follows that of Givens et al. (2009). Survey effort was concentrated in the fall between 140°W and 157°W longitude and south of 72° latitude. The total survey area was divided into 12 blocks as shown in Figure 1. Our analysis is limited to this region and to the period from August 28 to October 23. These limitations retain the vast majority of the fall BWASP data. A flight was deemed to be within the study area if no more than 10% of the plane positions recorded during the flight were outside the study area. This eliminated some spring flights and some flights over remote regions of the arctic outside the survey area, but no flights relevant for our analysis. Data for flights before 1982 were excluded due to differences in survey protocol and equipment. Data through 2006

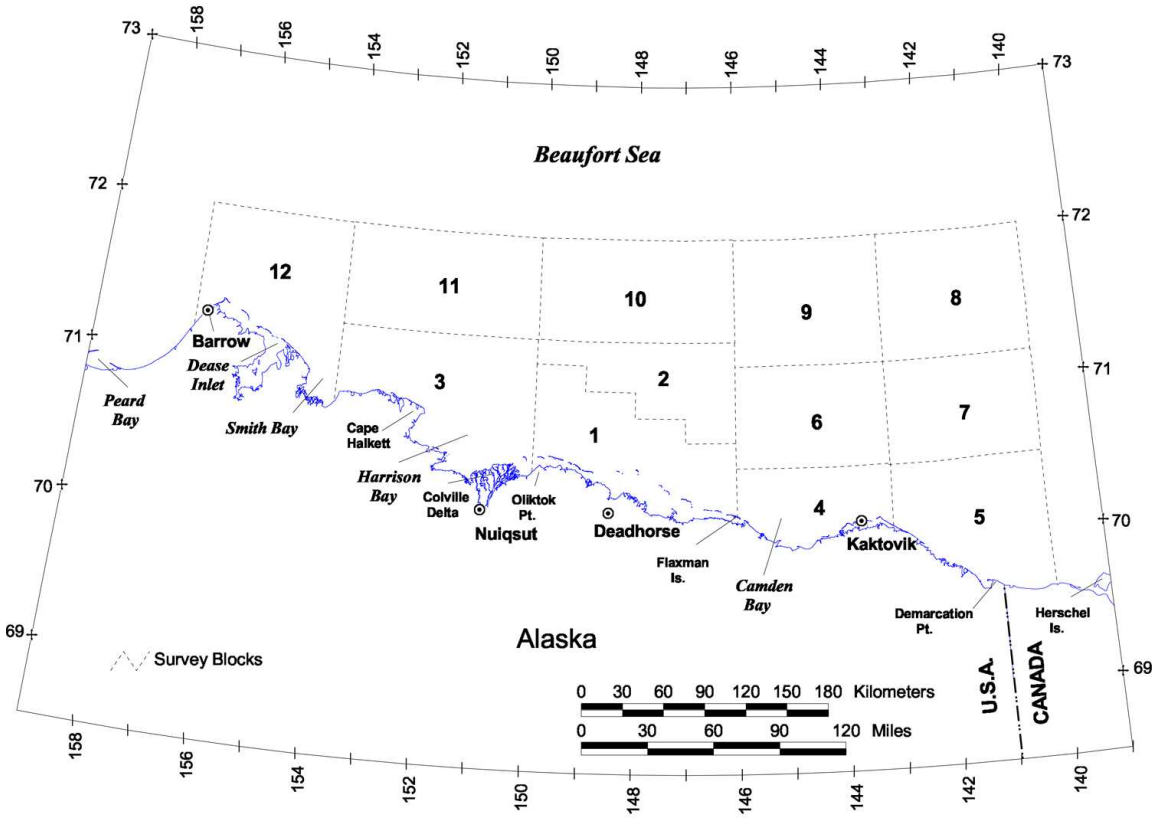

Figure 1: Map of BWASP survey blocks (Monnett and Treacy, 2005). Used with permission.

were available at the time of analysis. Figure 2 shows all sightings meeting these criteria.

Notwithstanding the visual impact of Figure 2, the BWASP data are in fact very sparse. If the survey region is divided into spatio-temporal blocks of 100km^2 $\times$ 1 week, the average number of blocks per sighting is 63, and that increases to 150 blocks when limited to the dataset we used (the 'on-transect' sightings; see below).

From 1982 through 2006, each year included between 23 and 93 flights. The pattern of block coverage was chosen somewhat opportunistically based on prevailing weather, a desire to investigate regions of potential industrial impact, and sometimes on suspected migration patterns. The latter choice violates one of the key assumptions of standard line transect analysis: that the transect lines should be randomly placed with respect to the distribution of animals. East of 154°W longitude, two-week spatio-temporal coverage was disproportionately targeted toward areas of suspected higher relative abundance. Thus, survey effort was relatively greater in regions near shore, thereby increasing total sightings by focusing on the primary migration corridor.

All flights began from Deadhorse, Alaska, near Prudhoe Bay. For a chosen survey block, a random transect grid was determined by dividing the block into 30-minute sections of longitude. Minute marks along both the northern and southern edges of each section were randomly chosen and connected with straight lines to create transect legs. This procedure was repeated for all 30-minute sections within the chosen block. Northern and southern

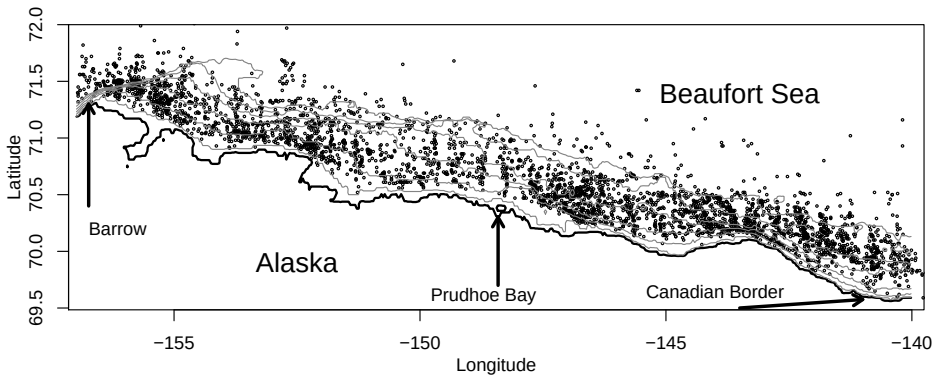

Figure 2: Map of all bowhead sightings within the spatio-temporal region used for the BWASP analysis. Each point indicates one cluster. Bathymetry contours from 5m to 65m in increments of 10m are shown with gray lines.

transect ends were connected alternately to form a flight path, with the first and last transect points connected to Deadhorse. One result of this design is that regions near the east and west boundaries of these 30-minute sections had comparatively less effort (M. Ferguson, US NMML, pers. commn). Flight deviations from the transect pattern were common due to flying conditions, sighting conditions, and other factors. Sighting distances were measured perpendicular to the transect line abeam of the plane and computed from altitude and clinometer readings. The nominal flight altitude was 458m.

The nature of survey effort varied during the flight. Each flight can be partitioned into discrete segments, with start and end points established by particular events during the flight. Each segment was labeled based on those events. The events that separated segments indicated initiation or termination of effort, sightings, and opportunistic records of survey conditions and covariates. The spatio-temporal location of every event was recorded. Flight segments to and from Deadhorse were denoted 'on-search', except that all flight portions over land were denoted as 'deadhead'. Segments wholly contained within a transect leg were denoted 'on-transect', with segments on connective legs between transects denoted 'on-connect'. Transect legs began and ended with 'start-transect' and 'end-transect' events. Normally sightings occurred at events labeled 'sighting-on-transect' or 'sighting-on-connect'. Occasionally, a possible cue or tentative sighting was detected. To investigate, a 'divert-transect' event was recorded and the flight continued with 'on-search' segments until a 'resume-transect'. During on-search effort, the goal was to confirm or refute that the possible sighting was a bowhead. Confirmation resulted in recording the sighting as occurring on-transect immediately before the divert-transect event. These on-search periods were generally characterized by a limited period of circling. If at any time during the on-search period additional whales were sighted, they were recorded as 'sightings-on-search'. Perpendicular distances to on-search sightings were measured in the same manner as other sightings, to the extent possible. Aside from rare missing data, every sighting was accompanied by measurement of covariate data such as sighting conditions. The covariate data were also recorded at convenient times between sightings. These events were labeled as 'point-on-transect',

'point-on-connect' and 'point-on-search'. These extra data present a complication beyond the simpler protocol where covariates are measured only at sightings. BWASP observers recorded sightings of animals besides bowheads, but we treated these as points on segments rather than sighting events. Rare circumstances required 'abort-flight' events.

The bowhead migration is age- and sex-structured during spring and fall (Moore and Reeves, 1993; Angliss et al., 1995). Also, at various places and times—such as for feeding aggregations—substantial clustering is observed. Although we have implied above that sightings correspond to animals, the reality is that sightings pertain to clusters. The number of animals in the cluster is one of the covariates recorded. Other covariates we examined included Beaufort sea state, ice coverage, day of sighting, and year of flight. Of course, the geographic coordinates of each event were also included in the BWASP dataset. Rather than use these directly, we re-expressed them as follows. We calculated the standardized distance of each event along an idealized shoreline linearly connecting the tip of Point Barrow to the shoreline at the Canadian border. Due to the nature of the shoreline along the north slope of Alaska, our shoreline distance measurements are quite similar to units of longitude. We replaced latitudinal positions with bathymetry data. Using a GIS database of bathymetry, each sighting was associated with the water depth at the sighting location (NOAA, 2008). Figure 2 shows contours of water depth from 5m to 65m in increments of 10m. Depth is strongly correlated with distance offshore, and analyses below indicate that depth has greater predictive power than distance off shore. There is a notable trough near Barrow (the Barrow Canyon) and the eastern Beaufort is generally shallower than the west. Bowhead feeding occurs mainly in the summer, in shallower water, and with greater clustering (Würsig and Clark, 1993; Griffiths and Buchanan, 1982; Wartzok et al., 1990). The predominant bathymetric zone of migration in the Beaufort Sea is 20-49m depth (Würsig and Clark, 1993). For analysis, sightings at depths exceeding 65m were excluded, eliminating a few extreme and rare observations with depths up to nearly 3000m.

2 Model Introduction

Now we focus on detailed and comprehensive modeling of the spatio-temporal relative density of bowheads and the relationship between density and various covariates. The approach here is built upon an underlying model introduced by Hedley and Buckland (2004). Hereafter we recognize their work as 'HB'. After reviewing their model, we later describe our approach.

Initially, assume that all data are from a single flight. For an aerial line transect survey with non-overlapping strips, parameterize the location of a point within the strip using a local coordinate system with x denoting the distance along the flight line from its origin to the orthogonal projection of the point onto the line, and y denoting the perpendicular distance away from the flight line (with negative values arbitrarily corresponding to the left side of the plane). The $|y|$ distances are truncated at a (possibly infinite) distance, w. Let $g(x, y)$ be the detection function, i.e., the probability of detecting an animal located at (x, y).

Let x_0 and x_{n+1} denote the locations of the start and end of survey effort, respectively, for a flight. Let $x_1, \ldots, x_n$ denote the locations of the n animals detected during the survey.

The starting location is given by $x_0 = 0$ and effort ceases as x_{n+1}. Defining $l_i = x_i - x_{i-1}$ for $i = 2, \ldots, n$, note that l_i denotes the "waiting distance" between sightings $i-1$ and i. Correspondingly, let l_1 and l_{n+1} denote waits until the first sighting and after the last.

Let $D(x, y)$ be the rate of an inhomogeneous Poisson process that governs the location of animals within the survey area. HB make the following assumptions about $g(x, y)$ and $D(x, y)$:

1. The Poisson intensity of animals is independent of the distance of the animal from the flight line: $D(x, y) = D(x)$. One reason this assumption is plausible is because the width of the survey strip is typically very small compared to the length of the track and total survey area, so wide-scale variation swamps any small variation across the strip.

2. The detection function is independent of the distance along the flight line and symmetric with respect to the distance off the track: $g(x, y) = g(y) = g(-y)$.

3. Detection of animals directly under the plane is perfect: $g(0) = 1$. Only in surveys with multiple observers or similar resighting opportunities can this assumption be relaxed.

4. Sighting distances are not correlated with location along the track: x and y are independent.

Letting $f_{l,y|x}$, $f_{l|x}$ and f_y denote the probability density functions of their respective subscripts, we can write the joint density of the observations as

$$f_{l,y|x}\left(\mathbf{l}, \mathbf{y}, |\mathbf{x}\right) = \left(\prod_{i=1}^{n} f_{l|x}\left(l_i|x_{i-1}\right)\right) \Pr_{L|x}\left[L_{n+1} > l_{n+1}|x_n\right] \prod_{i=1}^{n} f_y\left(y_i\right) \tag{1}$$

where $\mathbf{l} = (l_1, \ldots, l_{n+1})$, $\mathbf{y} = (y_1, \ldots, y_n)$, $\mathbf{x} = (x_0, \ldots, x_n)$, and the middle term above occurs because the last waiting distance is censored by the end of the flight.

The derivation of $f_{l|x}$ begins by noting that when the inhomogeneous Poisson process with rate $D(x, y)$ is thinned by an independent sighting process governed by $g(x, y)$, the locations of detections are realizations of a thinned inhomogeneous Poisson process with rate $D(x, y)g(x, y)$. Also, detections of a distance y off the trackline have probability density $f_y(y) = g(y)/2 \int_0^w g(y)dy$ (e.g., Buckland et al., 2001). Define the effective strip half-width to be $\mu_w = \int_0^w g(y)dy$. Then the cumulative distribution function $F_{l|x}$ can be expressed as

$$\begin{aligned}
F_{l|x}\left(l_i|x_{i-1}\right) &= 1 - \Pr\left[\text{no sightings in first } l_\text{i} \text{ distance beyond } x_{\text{i}-1}\right] \\
&= 1 - \exp\left\{-\int_{-w}^{w}\int_{x_{i-1}}^{x_{i-1}+l_i} D(x,y)g(x,y)\ dx\ dy\right\} \\
&= 1 - \exp\left\{-\int_{-w}^{w} g(y)\int_{x_{i-1}}^{x_{i-1}+l_i} D(x)\ dx\ dy\right\} \\
&= 1 - \exp\left\{-2\mu_w\int_{x_{i-1}}^{x_{i-1}+l_i} D(x)\ dx\right\}. \qquad (2)
\end{aligned}$$

For modeling spatial relative density, parameterize the spatial distribution and detection processes by vectors $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ respectively so $D(x) = D_\alpha(x)$ and $g(y) = g_\beta(y)$. Differentiation of (2) and substitution into (1) yields the likelihood

$$\begin{aligned} L\left(\boldsymbol{\alpha}, \boldsymbol{\beta} | \mathbf{l}, \mathbf{y}, \mathbf{x}\right) &= \left(\prod_{i=1}^{n} D_\alpha\left(x_i\right)\right) \times \\ & \exp\left\{-2\mu_w \sum_{i=1}^{n+1} \int_{x_{i-1}}^{x_i} D_\alpha(x) dx\right\} \prod_{i=1}^{n} g_\beta\left(y_i\right) \end{aligned} \tag{3}$$

after simplification.

Often the detection function depends on, say, J covariates observed along with y_i, namely $\mathbf{v}_i = (v_{i1}, \ldots, v_{iJ})$. In this case, the detection function can be replaced with $g_\beta(y_i|\mathbf{v}_i)$. The covariates $\mathbf{v}_i$ should comprise variables that affect aspects of the effective search effort. Since the BWASP protocol provides observations of covariates at locations aside from at x_i, the $\mathbf{v}_i$ should be considered data that are associated with the sighting attempt—not necessarily measured only upon sighting. The collection of all $\mathbf{v}_i$ in the dataset will be denoted $\underline{\mathbf{v}}$.

An over-simplification of this model would be that whale sightings occur along the flight lines according to a homogeneous Poisson process. Then the waiting distances between detections would be exponentially distributed. A compromise between the fully inhomogeneous model and simple homogeneity is explored by HB. They consider the assumption that the density of whales and the expected encounter rate are constant during each waiting period but may change at each sighting. With this assumption, each l_i has an exponential distribution with parameter λ_i.

The presence of a whale can depend on, say, K covariates. Denote these with the vector $\mathbf{z}_i = (z_{i1}, \ldots, z_{iK})$. For example, $\mathbf{z}_i$ may include ocean depth and day and year of flight. The collection of all $\mathbf{z}_i$ data will be denoted $\underline{\mathbf{z}}$.

Under the piecewise exponential model, relative density of whales in the strip surveyed during the waiting period can be estimated using a generalized additive model (GAM) with linear predictor expressed as

$$\eta^{-1}\left(E(l_i)\right) = \theta + \sum_k f_k\left(z_{ik}\right) \tag{4}$$

where $E(l_i) = 1/\lambda_i$, f_k are a collection of smooth functions estimated as part of the model fitting process, each of which expresses the effect of one spatial covariate z_k on relative density. Here, η^{-1} is the link function; for the above exponential distribution, the log link $\eta^{-1}(1/\lambda) = \log(1/\lambda)$ would be appropriate.

Let $\hat{l}_i$ denote the fitted wait for the ith sighting. Then an index of relative abundance at sites (x_i, y_i) is proportional to $1/\hat{l}_i$.

2.1 Expanded Model

Application to the BWASP dataset requires some revision and extension of the above approach. Topics meriting further consideration include (i) censoring, (ii) clustering, (iii) co-

variate effects for both the presence and detection processes, and (iv) covariate measurement.

2.1.1 Censoring

With respect to the above likelihood function (3), we view censoring a bit differently. Clearly the trailing flight segment after the last sighting yields a (right) censored waiting distance, l_{n+1}. In our view, the waiting distance between the first instance of effort for the flight and the first sighting is also a right-censored observation because its duration is known to exceed l_1. In a classic right-censoring situation, a treatment is applied at the initiation of observation effort and outcomes are censored if observation effort ceases before the outcome is observed. In the present case, there is no relevance to the direction of time because no treatment begins when the observation effort begins. Censoring occurs when the timespan of observation effort fails to cover both endpoints of an inter-sighting distance and we therefore know only a lower bound for the time between sightings.

Furthermore, the BWASP dataset includes many flights, and many of these yielded no sightings ('null flights'). These null flights provide information about relative density nevertheless. Since flights were made on different days, it is inappropriate to string together waiting distances across days due to changes in covariates. Thus a null flight should also be treated as providing a censored waiting distance known to exceed the total effort distance on that flight.

Let $a_i = 0$ if the ith waiting period is censored and $a_i = 1$ if it is not. Then accounting for all censoring, the joint density of the observations for any single flight given in (1) becomes

$$\begin{aligned} f_{l,y|x}\left(\mathbf{l},\mathbf{y}|\mathbf{x}\right) &= \left(\prod_{i=1}^{n+1} f_{l|x}\left(l_i|x_{i-1}\right)^{a_i} \times \Pr_{L|x}\left[L_i > l_i|x_{i-1}\right]^{1-a_i}\right) \times \\ & \prod_{i=1}^{n} f_y\left(y_i\right) \end{aligned} \tag{5}$$

and the overall density is the product of such terms over all flights. Aside from the introduction of multiple flights, this revision is inconsequential in (3) because of the form of $F_{l|x}$, however it introduces an important censoring component in our estimation approach described later.

2.1.2 Clustering

Another important aspect of the BWASP data is that the sightings represent detections of whale clusters, not individual whales. Clusters may contain one or more whales. Cluster size can influence detection probabilities. Moreover, clustering varies with space and time for bowheads due to the nature of their migration and behavioral patterns. Aside from cow-calf pairs, bowheads tend to cluster in the times and places associated with breeding and feeding (**ref**).

To model clusters, we introduce a random variable C_i whose observed value c_i is the size of the cluster observed at (x_i, y_i). To the previous assumptions we add the assumption that

waiting period and distance off trackline are conditionally independent given cluster size. Then $f(l, y, c|x) = f_{l|c,x}(l|c, x) f_{y|c}(y|c) f_{c|x}(c|x)$ since $y|c$ is independent of x.

We also introduce the censoring indicator b_i which equals 0 when c_i is missing (in which case only the existence of one or more whales is known) and 1 when it is not. Then for $i = 1, \ldots, n+1$, $a_i b_i = 1$ indicates a completely uncensored observation (l_i, c_i), while $(1-a_i)b_i = 1$ indicates lead censoring and $(1-a_i)(1-b_i) = 1$ indicates trailing censoring and null flights. Now the joint density of the observations becomes

$$\begin{aligned} f_{l,y,c|x}\left(\mathbf{l}, \mathbf{y}, \mathbf{c}|\mathbf{x}\right) &= \\ &\left[\prod_{i=1}^{n+1} f_{l|c,x}\left(l_i|c_i, x_{i-1}\right)^{a_i b_i} \times \Pr_{L|c,x}\left[L_i > l_i|c_i, x_{i-1}\right]^{(1-a_i)b_i} \times \right. \\ &\left. \left(\sum_c \Pr_{L|c,x}\left[L_i > l_i|c, x_{i-1}\right] f_{c|x}\left(c|x_{i-1}\right)\right)^{(1-a_i)(1-b_i)}\right] \times \\ &\prod_{i=1}^{n} f_{y|c}\left(y_i|c_i\right) \prod_{i=1}^{n} f_{c|x}\left(c_i|x_i\right). \end{aligned} \tag{6}$$

The term in parentheses marginalizes out c for trailing censoring and null flights, where cluster size cannot be observed.

2.1.3 Covariate effects for presence and detection processes

Introducing covariate effects $\mathbf{v}$ into the detection function as $g(y|c, \mathbf{v})$ is very helpful for the BWASP data because sighting conditions such as sea state and ice coverage vary greatly across the region and during the survey season. Similarly, covariates are also important for modeling the presence process since relative abundance depends on space, time, and other factors. For the wait l_i, the presence covariates are denoted $\mathbf{z}_i = (z_{i1}, \ldots, z_{iK})$. The $\mathbf{z}_i$ should be considered to represent data associated with the ith wait, which is not to imply that they must be observed at x_i. One reason for this distinction is that the BWASP protocol provides covariate observations at many distinct locations between sightings (see Section 1). For notational simplicity, we incorporate x_{i-1} into $\mathbf{z}_i$. Thus we now write the Poisson delivery rate as $D(c, \mathbf{z})$. Adapting for all covariate effects, we have the joint density

$$\begin{aligned} f_{l,y,c|z,v}\left(\mathbf{l}, \mathbf{y}, \mathbf{c}|\underline{\mathbf{z}}, \underline{\mathbf{v}}\right) &= \\ &\left[\prod_{i=1}^{n+1} f_{l|v,z}\left(l_i|c_i, \mathbf{z}_i, \mathbf{v}_i\right)^{a_i b_i} \times \Pr_{L|c,v,z}\left[L_i > l_i|c_i, \mathbf{v}_i, \mathbf{z}_i\right]^{(1-a_i)b_i} \times \right. \\ &\left. \left(\sum_c \Pr_{L|c,v,z}\left[L_i > l_i|c, \mathbf{v}_i, \mathbf{z}_i\right] f_{c|z}\left(c|\mathbf{z}_i\right)\right)^{(1-a_i)(1-b_i)}\right] \times \\ &\prod_{i=1}^{n} f_{y|c,v}\left(y_i|c_i, \mathbf{v}_i\right) \prod_{i=1}^{n} f_{c|z}\left(c_i|\mathbf{z}_i\right). \end{aligned} \tag{7}$$

It may be unclear initially why the conditional distribution of l_i is dependent on $\mathbf{v}_i$ above. Note that a derivation akin to (2) yields

$$F_{l|c,v,z}\left(l_i|c_i,\mathbf{v}_i,\mathbf{z}_i\right) = 1 - \exp\left(-2\mu_w(c_i,\mathbf{v}_i)\int_{x_{i-1}}^{x_{i-1}+l_i} D\left(x,c_i,\mathbf{z}_i\right)dx\right) \tag{8}$$

where $\mu_w(c_i,\mathbf{v}_i) = \int_0^w g(y,c_i,\mathbf{v}_i)dy$. Using the parameterization like (3) with $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, and $\boldsymbol{\gamma}$ parameterizing the presence, sighting, and cluster processes respectively, we obtain the likelihood

$$\begin{aligned}
&L\left(\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}|\mathbf{c},\mathbf{l},\mathbf{y},\right) \quad = \\
&\left\{\prod_{i=1}^{n+1}\left[D_\alpha(x_{i-1}+l_i,c_i,\mathbf{z}_i)\left(1-F_{l|c,v,z}(l_i|c_i,\mathbf{v}_i,\mathbf{z}_i;\boldsymbol{\alpha},\boldsymbol{\beta})\right)\right]^{a_ib_i}\times\right.\\
&\left(1-F_{l|c,v,z}(l_i|c_i,\mathbf{v}_i,\mathbf{z}_i;\boldsymbol{\alpha},\boldsymbol{\beta})\right)^{(1-a_i)b_i}\times\\
&\left.\left(\sum_c\left(1-F_{l|c,v,z}(l_i|c_i,\mathbf{v}_i,\mathbf{z}_i;\boldsymbol{\alpha},\boldsymbol{\beta})\right)f_{c|z}(c|,\mathbf{z}_i;\boldsymbol{\gamma})\right)^{(1-a_i)(1-b_i)}\right\}\times\\
&\prod_{i=1}^{n}g_\beta(y_i|c_i,\mathbf{v}_i)f_{c|z}(c_i|\mathbf{z}_i;\boldsymbol{\gamma})
\end{aligned} \tag{9}$$

There are two important features of this likelihood. First, it is important to note that the only way that $L(\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma})$ depends on $\boldsymbol{\beta}$ is through the effective strip half-width, $\mu_w(c_i,v_i,\boldsymbol{\beta})$. Thus we may write the likelihood as $L(\boldsymbol{\alpha},\mu_w(c_i,v_i,\boldsymbol{\beta}),\boldsymbol{\gamma})$. Second, the likelihood involves clustering in two ways: the distribution of cluster sizes can depend on spatial covariates, and detection probabilities can depend on cluster size.

In summary, the likelihood is set up so that $\boldsymbol{\alpha}$ parameterizes the dependence of waiting periods for clusters (l) on cluster sizes (c) and covariates that affect presence ($\mathbf{z}$). The dependence of sighting distances (y) on cluster sizes and sighting covariates ($\mathbf{v}$) is parameterized by $\boldsymbol{\beta}$. The dependence of cluster sizes on presence covariates is parameterized by $\boldsymbol{\gamma}$.

2.1.4 Covariate measurement

Finally, note that for BWASP there are (usually) more than one set of covariate observations associated with a waiting distance. These correspond to a variety of intra-wait segments such as those created by points-on-transect. To associate covariate values with a wait, we take the approach of using a weighted average of observations during the wait. The weight for each segment is taken to be proportional to the length of the segment that leads to the recording point. Segments with missing covariate values at the endpoint are ignored.

This approach is used for deriving covariates to be used in the analysis of waiting distances because the observed wait reflects whale density flown over during the wait. We use a different approach for deriving covariates to used in the estimation of detection probabilities. In this

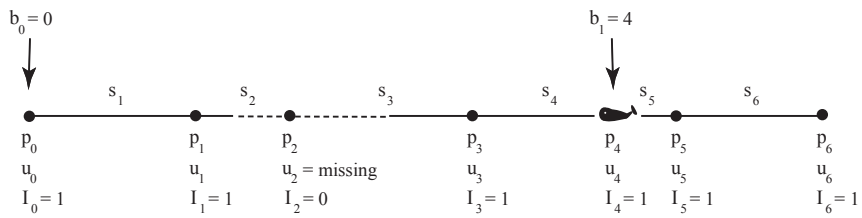

Figure 3: Illustration of centered segment weighted averaging. At point p_4 a bowhead was sighted. The covariate values u_1, u_3 and u_4 contribute (in proportion to their respective surrounding segment lengths) to the average covariate value during the wait until the sighting (at the fourth point here). The value at u_2 cannot contribute because it is missing.

case, we are uncomfortable with the piecewise homogeneous approach that assigns covariate values from observations only at the initiation of a new waiting period. This strategy is appropriate when there are no inter-sighting data and waiting distances are short compared to variation in sighting conditions. For the present case this approach ignores most BWASP covariate data. Furthermore, over long waiting periods the leading covariate observation could be an poor summary of conditions at the point of detection. The detection function describes the probability of sighting a whale (cluster) at a specific distance, given covariates. The covariates that pertain are those that modify this relationship between distance and detection. To address this, we retain the assumption that the expected detection rate is constant over the waiting period but modify the covariate calculation. Again we take a weighted averaging approach, but the relevant segments are taken to be among those that surround the sighting. We denote this approach as centered weighting.

Specifically, for centered weighting covariate averages were computed as follows (e.g., Figure 3. Each waiting distance comprises a collection of one or more shorter on-effort segments determined by records of point-on-transect, etc., as described in Section 1. A covariate average of a waiting distance is defined to be the segment length-weighted average of the covariate values over all non-missing data during the wait. For a given flight, let M denote the number of segments, determined by records at points $p_0, \ldots, p_M$. Denote the between-point segment lengths as $s_1, \ldots, s_M$. Let the covariate u be measured at some or all of the points, where $u_m = 0$ if u is not observed at p_m. Next, let the indicator I_m equal 1 if u_m is observed and 0 otherwise. I_m will be used to ensure that u_m contributes to the average only when $I_m = 1$. Finally, let integers $b_j \in \{1, \ldots, M\}$ index sightings and locations so that $b_j = k$ if the jth sighting occurs at p_k. The weighted average covariate value corresponding to the jth sighting is given by

$$\bar{u}_{b_j} = \frac{I_{b_{j-1}} u_{b_{j-1}} s_{b_{j-1}+1} + \sum_{m=b_{j-1}+1}^{b_j} I_m u_m (s_m + s_{m+1})}{I_{b_{j-1}} s_{b_{j-1}+1} + \sum_{m=b_{j-1}+1}^{b_j} I_m (s_m + s_{m+1})} \tag{10}$$

for $j = 1, \ldots, B$ where the total number of sightings on the flight is $b \geq 1$, $b_0 = 0$, and s_{b_j+1} always exists since no flight ends with a sighting. Null flights are handled analogously. Actually, the above description over-simplifies the definition of I_m: there are reasons aside from missing data when u_m should not contribute during the waiting period. For example, any single segment connecting deadhead to abort-flight should not count in the average.

2.1.5 Multiple flights

The BWASP data include between 23 and 93 flights per year over nearly thirty years. The overall likelihood, therefore, is taken to be product of terms like (9) over all flights.

2.2 Model Fitting

Like the original HB likelihood, our expanded version is complex and would be difficult to maximize directly in most situations. However, some parametric assumptions and estimation approaches can be adopted to facilitate approximate inference about relative density.

In the generic model described above, the Poisson process is inhomogeneous. A compromise between the fully inhomogeneous model and simple homogeneity is explored by HB. They consider the assumption that the density of whales and the expected encounter rate are constant during each waiting period but may change at each sighting. Under this model, each l_i has an exponential distribution with a (potentially) unique mean.

This approach can be adapted for our model and its application to the BWASP data. Assume that cluster size is observed any time that l_i is uncensored, and adopt the model that $\mathrm{E}(l_i) = \exp\{-\psi_i\}$. This implies that relative animal density is high when ψ_i is high. It also suggests using a generalized additive model (GAM) for waiting distances given $\mathbf{c}$ and $\mathbf{y}$, namely

$$\psi_i = \log \lambda_0 + g_c(c_i) + \sum_j g_{vj}(v_{ij}) + \sum_k g_{zk}(z_{ik}) + \tag{11}$$

A log link function is inherent in (11), and the univariate smooth functions g whose subscripts correspond to the covariates describe how the variables affect l_i. This exponential model effectively replaces $f_{l|c,y}(l_i|c_i, y_i; \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma})$ with an alternative parameterization $f_{l|c,y}(l_i|c_i, y_i; \psi_i)$. The dependence of waiting periods on the presence and sighting of clusters is captured by g_c. For BWASP, we parameterize g_c as $g_c(c|\boldsymbol{\eta})$ which is defined to equal η_1 if $c = 1$, η_2 if $c = 2$, and η_3 if $c \geq 3$. We have expressed the other presence effects with smooth functions g_{zk} and sighting effects with g_{vj}. Of course, the univariate functions g could be replaced or supplemented with multi-way interaction terms, i.e., smooth functions of several covariates.

For this GAM, the likelihood is proportional to the product over all flights in all years of terms like

$$\prod_{i=1}^{n+1} f_{c|z}(\boldsymbol{\gamma}|c_i, \mathbf{z}_i)^{a_i} f_{y|c,v}(\boldsymbol{\beta}|c_i, \mathbf{v}_i, y_i)^{a_i} \left[(e^{\psi_i} \exp\{-l_i e^{\psi_i}\})^{a_i} (\exp\{-l_i e^{\psi_i}\})^{1-a_i} \right]. \tag{12}$$

We estimate ψ using the marginal likelihood arising from the distribution of $\mathbf{l}|\mathbf{c}, \mathbf{y}$. From our exponential GAM model, then, inference is based on

$$\log L = \sum_{i=1}^{n+1} \{a_i(\log f(l_i|\psi_i) + (1 - a_i) \log S(l_i|\psi_i)\} \tag{13}$$

where f and S are the density and survivor functions parameterized by ψ and given in the square-bracketed portion of (12). This can be simplified as

$$\begin{aligned}
\log L &= \sum_{i=1}^{n+1} \{a_i(\psi_i - l_i \exp\{\psi_i\}) + (1-a_i)(-l_i \exp\{\psi_i\}\} \\
&= \sum_{i=1}^{n+1} a_i(\psi_i + \log l_i) - l_i \exp\{\psi_i\} - a_i \log l_i \\
&= Q + \sum_{i=1}^{n+1} a_i \log(l_i \exp\{\psi_i\}) - l_i \exp\{\psi_i\} \qquad (14)
\end{aligned}$$

which follows because $\sum_i a_i l_i$ is independent of the components of ψ and Q is a constant. Note that (14) is identical to the log likelihood for the model where $a_i \sim \text{Pois}(l_i \exp\{\psi_i\})$. This reflects a well-known result in generalized linear modeling for censored exponential data (McCullagh and Nelder, 1989). Therefore we may estimate the effects expressed by the terms in ψ using a GAM for a Poisson distribution with mean a_i and using the log link, namely

$$\log \mathrm{E}(a_i) = \log l_i + \psi_i. \qquad (15)$$

In the GAM framework, $\log l_i$ is an offset and the components of ψ can be estimated using standard GAM fitting methods Wood (2000; 2004). From the estimates $\hat{\psi}_i$, we can easily obtain fitted values $\hat{l}_i$ corresponding to the exponential version of the model since (13) is identical to the Poisson likelihood in (15). Specifically, $\hat{l}_i = \exp\{-\hat{\psi}_i\}$.

Estimation of sighting probabilities, and hence $\hat{\mu}_w$ values, can be obtained by applying the standard DISTANCE approach for estimation of a line transect detection function (Thomas et al., 2006; Buckland et al., 2001, 2004). Finally, $\boldsymbol{\gamma}$ are nuisance parameters whose values are uninteresting for the present purpose because. Thus, given values for the cluster size c_i and for the other covariates, the overall relative density at (x_i, y_i) can be estimated as

$$R_i = c_i / \hat{l}_i \hat{\mu}_w. \qquad (16)$$

Because of the potentially strong dependencies within a dataset like BWASP, asymptotic tests of statistical significance for model terms are not appropriate. Furthermore, Wood (2006) notes that p-values will often be underestimates if smoothing parameters have been estimated as part of the model fitting. We turn to bootstrapping instead, beginning by resampling years with replacement. For each instance of a resampled year, flights within year are resampled with replacement. Finally, for each resampled flight, waiting distances are resampled with replacement. All bootstrap sampling sizes matched the original data. This approach preserves some potential correlation since both years and flights are sampling blocks. It does not divide observations within a flight into blocks. The scarcity of sightings, censoring, and null flights would complicate such fine-scale blocking.

3 Application to BWASP

3.1 Model fitting

For the BWASP data, we chose to use only sightings and censored waits recorded on-transect because the nature of on-connect and on-search effort and data recording were comparatively poorly understood. GAMs were fit using the approach of Wood (2000; 2004). Model choice was conducted using forward selection based on AIC values (e.g., Burnham and Anderson, 1998) with an improvement of 2 AIC units required for model expansion. A collection of possible model terms was identified to ensure that the final model would adequately address the nature of the fall bowhead migration using key covariates and spatio-temporal dependencies. The covariates allowed for terms that might contribute to the final model for waiting distances were cluster size, day, year, ocean depth, and distance along idealized shoreline.

Modeling the effects of these variables relied foremost on fitting additive univariate and bivariate tensor product splines as described by Wood (2007). This approach allows modeling the spatio-temporal pattern of the fall migration which is known to proceed westward with potential variation in the distance of whales offshore (and therefore variation in depth). Higher order terms were excluded for simplicity. Appropriate degrees of freedom for spline terms were simultaneously estimated using the unbiased risk estimator criterion (UBRE) of Wood (2000; 2004).

Year was expressed as the number of years since 1982. This term can compensate for known increases in bowhead population size (George et al., 2004) and possible smooth inter-annual variation in migratory behavior associated with variation in environmental or anthropogenic factors. Abrupt inter-annual shifts are not modeled here.

Effects for cluster size were modeled as an additive effect associated with a categorical variable having three levels: 1 animal, 2 animals or at least 3 animals. This choice was motivated by an extremely skewed distribution of cluster sizes, and by the belief that the relationship (if any) between spatio-temporal patterns of whale presence and cluster size is highly non-linear in cluster size. Interactions between cluster size and each other variable were modeled by allowing three separate splines for the other variable–one per cluster level. During a preliminary exploratory phase, effects for cluster size were also examined using a variety of other data transformations and using the raw counts. These choices were inferior.

Table 1 shows the resulting sequence of models. At no point was there convincing evidence that inclusion of cluster size improved the model. This indicates that the arrival rate of sightings or the inter-sighting waiting distances do not depend on cluster size after controlling for the other spatio-temporal variables. This result is somewhat counterintuitive. Waits between clusters are driven by two factors: the spatial distribution of clusters and the relative detection probabilities for clusters of varying sizes. Apparently, the former variation is adequately explained by how the covariates affect waiting distances. The latter factor is discussed below.

In addition to GAM fitting, estimation of relative density relies on corrections for sighting

Model	AIC
t(as,dp)+t(as,yr)+t(dy,dp)+t(yr,dp)+t(as,dy)+ t(yr,dy)+t(dp)	4493.19
t(as,dp)+t(as,yr)+t(dy,dp)+t(yr,dp)+t(as,dy)+ t(yr,dy)	4498.96
t(as,dp)+t(as,yr)+t(dy,dp)+t(yr,dp)+t(as,dy)	4534.31
t(as,dp)+t(as,yr)+t(dy,dp)+t(yr,dp)	4588.46
t(as,dp)+t(as,yr)+t(dy,dp)	4633.10
t(as,dp)+t(as,yr)	4755.18
t(as,dp)	5021.06

Table 1: AIC values for a sequence of models. Abbreviations are: tensor spline (t), day (dy), depth (dp), distance along idealized shoreline (as), and year (yr).

probabilities, as represented by the estimated effective strip half-widths $\hat{\mu}_w$ at each sighting. These corrections were estimated using the method of Thomas et al. (2006) and Buckland et al. (2001; 2004), which fits a parametric model for detection probabilities as a function of distance off trackline and covariates related to sighting ease. We used the hazard function model with no expansion terms. For this estimation only, observations exceeding 5280ft were excluded to improve model robustness (Buckland et al., 2001). Covariates tested for their effect on detection probability (in addition to distance off trackline) were: Beaufort sea state, cluster size, ice coverage, and visibility rating. These variables are explained further by Treacy (2002) and Monnett and Treacy (2005).

Of the covariates related to detection, only Beaufort sea state significantly altered the relationship between sighting distances and detection probabilities. Beaufort sea state is a measure of sea surface action, ranging from glassy to huge waves in hurricane-force winds. The predominant sea states in the BWASP data are states 1 through 4 (ripples through small waves), with the mode being state 2 (small wavelets). Very few high sea states were recorded. We categorized Beaufort sea state into three levels: low (glassy and light ripples), moderate (small and large wavelets), and high (small waves and worse).

Contrary to expectations, cluster size did not affect the detection function. This may reflect the fact that most sightings occurred in good sea states and good visibility conditions with little sea ice. Also, the distribution of cluster sizes is highly positively skewed. There may be too few data to obtain a reliable estimate of cluster size effects, particularly since the comparative impact of large cluster sizes is likely most severe at extreme distances but sighting distances were truncated at 5280ft for analysis. Despite the absence of a significant cluster size effect in both the GAM model and this detection function estimation, final estimates of relative density are still adjusted for cluster size using the estimator in (16).

Figures 4 and 5 shows maps of relative whale densities for six days in each of four years. Darker tones correspond to higher density, and a common relative density scale is used throughout all 24 maps. Contours show deciles within each image. The axes in each plot show the distance along the idealized shoreline and the water depth. However, it is important to note that the relationship between depth and distance to shoreline is distorted on the left

edge of these images because near shore depths are deeper due to an oceanographic feature known as the Barrow Canyon. The aspect ratio for these panels was chosen to resemble actual great circle distances. Together, these maps span most of the temporal and spatial region of data. The relative densities shown here represent the estimates from the GAM after full correction for detection probabilities and cluster sizes, as given by (16).

3.2 Relative density estimates

The space-time maps shown in Figures 4 and 5 provide a good representation of the conventional wisdom about bowhead migration while providing substantial new detail. Generally, they show an abundance of animals in the northeast Beaufort Sea in early fall. Particularly in more recent years, some of these whales move into the southeast Beaufort to shallower waters as they begin to move westward. The main migration then begins with a large number of whales traveling westward toward Barrow in 30-50m of water. Some whales linger in the east but swim west in smaller numbers as the season progresses. Even in late October some bowheads may not have headed west. Note that the westward migration cannot be characterized as a steady stream of whales. Instead, there appears to be a large initial pulse with numerous dispersed stragglers.

These figures and our results for other years also show substantial iterannual variation, particularly in the daily westward progress. This variation appears to exhibit both systematic trends and random fluctuations. Generally, it appears that the migration pulse has become earlier over the years.

Across the region, densities have increased markedly over the last 20 years. This is consistent with a time series of ice-based visual and acoustic census data indicating steadily increasing abundance over the same period (George et al., 2004).

Our maps also provide evidence that a group of whales is already present near Barrow in late August and that the relative abundance of such whales has increased in recent years. Although this could reflect an increasingly early component of the migration, it is also consistent with hunter observations of apparent 'summer resident' bowheads off Barrow who do not travel any further east, summer near Barrow, and rejoin the main migration as it returns in the late fall. Hunters also report seeing bowheads linger in the east much later in the season in recent years. Our results are consistent with these reports as well.

3.2.1 Localized relative density variation

Hypothesis motivation Perhaps the most striking aspect of the maps in Figures 4 and 5 is seen north of Prudhoe Bay (and to a lesser extent somewhat west of this). First, there is an unusual, persistent region of low local abundance (a 'hole') in waters less than 30m deep in the central coastal region. At these longitudes the whales show a strong tendency to prefer deeper waters (40-50m). Even though the shape and depth of this density 'hole' vary interannually, the general location does not.

The relative scarcity of whales in the 'hole' can also be seen in the raw data shown in Figure 2. Examination of this graph led some scientists to hypothesize that this region rep-

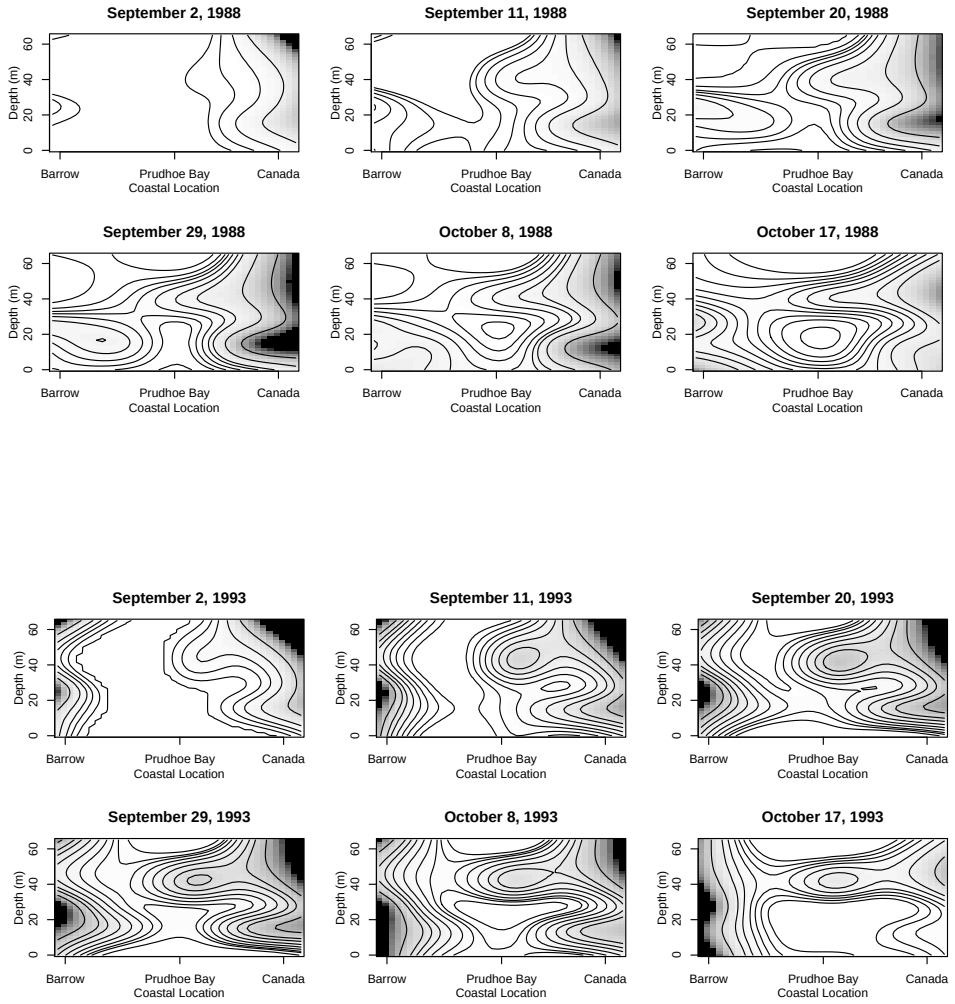

Figure 4: Fitted relative whale densities for six dates in 1988 and 1993. Higher densities correspond to dark colors, and all images in Figures 4 and 5 are on a common scale. Contours represent deciles within each image.

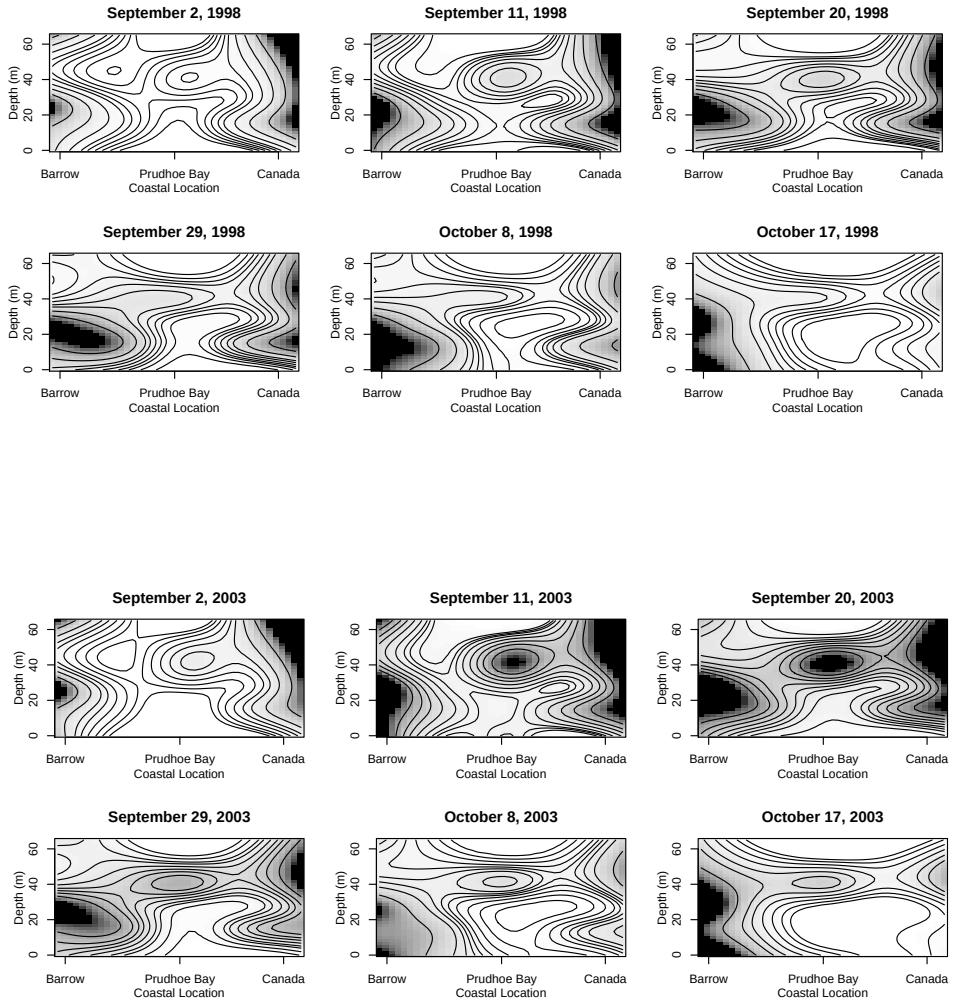

Figure 5: Fitted relative whale densities for six dates in 1998 and 2003. Higher densities correspond to dark colors, and all images in Figures 4 and 5 are on a common scale. Contours represent deciles within each image.

resents an area of lower relative density. However, when the data are limited to on-transect records, the 'hole' is much less visually compelling. This casts some doubt about the reliability of informal observations based on all sightings without correction for effort and sighting type. Statistical techniques can provide an objective empirical basis for consideration of a 'hole'.

If the apparent 'hole' represents a region of lower relative abundance associated with anthropogenic factors, it is potentially very important because disturbances to the whale migration can have consequences for the health and well-being of the whale population, as well as for local villagers who rely on their subsistence harvest. If the 'hole' is related to environmental factors, it is important to identify them and consider their import for whale and ecosystem management.

There are many possible hypotheses about the 'hole' than could be entertained. Industrial activities (especially acoustical exploration) might induce spatial deflections such as whales moving further offshore. There are other plausible explanations for regions of low relative abundance including environmental factors like prey availability and anthropogenic factors like aboriginal hunting activity. None of these factors have well-understood impacts, and the relative plausibility of the corresponding hypotheses is uncertain.

It is also important to realize that regions like the 'hole' need not reflect lower whale density. The apparent scarcity of whales may reflect aspects of availability, namely the proportion of time during which a whale is able to be detected by the aerial surveyor. In a region like the 'hole', whales may be swimming faster or be spending a greater proportion of time underwater.

Testing To investigate the apparent 'hole', we created several additional covariates. The first variable ('hole proximity') was the great circle distance from the sighting location to the center of the purported 'hole' (148.05°W, 70.63°N), as identified by C. Monnett (US MMS, pers. commn). A more localized proximity variable ('local hole proximity') was created by re-expressing this proximity variable on a scale that provided exponential decay in proximity, thereby emphasizing the local nature of the 'hole'. Discrete indicators of the 'hole' were also used. First, we examined a binary variable ('small box') that indicated whether the sighting was in a box encompassing the putative 'hole'. The box was bounded by 148.70°W, 147.90°W, 70.48°N and 70.74°N. Next, we expanded the region ('big box') to indicate if the sighting was between 147.90° W and 151.06° W, with no latitudinal boundaries. Based on typical swim speeds (Zeh et al., 1993), a bowhead would require only about one day to traverse the 'big box' at a consistent medium to fast swim speed.

If the 'hole' represents a disturbance reaction by the animals, one might expect to find milling, 'lingering' or'piling up' behavior, i.e., increased local density, upstream (east) of the 'hole'. Downstream (west) of the 'hole', one might expect to find a relative scarcity of animals for some distance until the effect of the disturbance has waned. To test this scenario, we created a contrast variable ('hole contrast') which equaled -1 for sightings between 146.25°W and 147.90°W, 1 between 147.90°W and 151.06°W, and 0 everywhere else.

Some of the motivation for studying 'hole' comes from scientists' previous informal reviews of the BWASP dataset itself. An industry impact hypothesis formed wholly independently from the dataset can be based upon the development of BP Exploration Inc.'s Northstar project built on the submerged remnants of Seal Island a bit northwest of Prudhoe Bay and about 54 miles northeast of Nuiqsut. Construction and subsequent operation of this facility began in 2001. We created a variable ('Northstar proximity') that measured sighting distance from Seal Island since Northstar's construction. This variable was trimmed at 37km distance because previous study of Northstar noise had indicated that 37km was roughly the most distant limit where directional autonomous sea floor acoustic recorders and sonobuoys detected noise associated with Northstar (LGL, 2004). The noises detected at greatest distances were generally vessel noises. Finally, a binary variable ('Northstar disk') indicated whether the sighting was within this 37km range. Sample sizes for these variables are very small.

Results The seven hypotheses represented by these variables are listed in Table 2. Each of these seven variables was tested separately in the final model with p-values estimated using our bootstrap approach. Table 2 also provides comments and concerns about each test.

For the 'hole', we found that 'hole proximity', 'local hole proximity', and 'big box' provided significant effects representing scarcity near the 'hole', with p-values of $p < 0.001$, $p = 0.003$ and $p < 0.001$ respectively. The 'small box' variable did not show a significant negative effect. The effect of the 'hole contrast' was also significant ($p < 0.001$), with 'piling up' in the east and scarcity in the west. For the variables related to Northstar, we found mixed results. The 'Northstar proximity' variable did not have a significant effect on relative density. However, the 'Northstar disk' variable was significant ($p = 0.001$), showing scarcity in the region.

Discussion Taken together, these results suggest that there are indeed lower apparent relative whale densities in the tested regions. It is worth noting that the tested regions are located near the middle of the data region, where estimation variability should be least, accuracy should be best, and extrapolation is nonexistent.

However, with the exception of 'big box' and 'hole contrast', these variables are assessed on a much finer scale than inter-sighting distances. In these cases, sample sizes are very small. Caution is warranted when trying to infer small-scale effects from a small number of observations obtained from a survey that is implicitly larger-scale. Therefore, aside from the 'hole contrast' and 'big box' comparisons, our results are not wholly persuasive.

The significance of the 'hole contrast' variable is particularly notable because it provides evidence that whales are 'piling up' before crossing the 'hole'. This sort of behavior would seem to be more easily explained as an avoidance reaction than as an effect of prey availability or migratory route preference. To gain confidence in the avoidance hypothesis, it is critical to examine whether the 'piling up' result could be caused, for example, by higher prey abundance or behavioral tendencies (e.g., more milling) in that region.

Hypothesis	P-value	Comments
'small box'	0.900	Small sample size; data-driven hyp.
'hole proximity'	<0.001	Same as 'small box'.
'local hole proximity'	0.003	Same as 'small box'.
'big box'	<0.001	Partially data-driven hyp.
'hole contrast'	<0.001	Same as 'big box'.
'Northstar proximity'	0.237	Very few points with high proximity.
'Northstar disk'	0.001	Very small sample size.

Table 2: Outcomes of hypothesis tests, with associated comments. In each case, the alternative hypothesis is that relative abundance decreases in the region or for increased proximity to it.

	Behavior					Swim Speed		
Region	swim	mill/rest	play	feed	other	slow	medium	fast
East contrast	78	15	3	1	4	31	64	5
West contrast	77	14	5	2	2	26	61	12
Elsewhere	78	12	4	3	3	33	60	7

Table 3: Percentages of behaviors, and separately of swim speeds, for whales in the two portions of the contrast region and elsewhere.

For some sighted animals, behavior was recorded. We classified behaviors into swimming, milling/resting, feeding, playing, and other. Table 3 shows the frequencies of these behaviors for on-transect sightings. These frequencies are calculated for three spatial areas: the western contrast area, the eastern contrast area, and everywhere else. There is no significant difference in behavior between these three regions. This provides no support for the hypothesis that the contrast finding is related to behavioral differences like milling tendencies.

Another potential explanation is that the regions of low animal density are related to increased swim speeds. We examined records of estimated swim speed categories (> 4km/hr, 2-4km/hr, and 0-2km/hr) for the portion of swimming whales where such data were available. The observed frequencies are shown in Table 3 for the three regions. Although there is no statistically significant difference in swim speeds (p=0.06), Table 3 shows that there are a few more fast swimmers in the western contrast block than the eastern one. Such a trend is consistent with a hypothesis that 'piling up' is followed by a period of quicker westward migration over the next 120km (approximately one day at a consistent medium to high swim speed).

3.2.2 Temporal trends in localized density variation

Finally, we can ask whether a zone of localized scarcity has changed over time. This could be informative, for example, if a 'hole' is attributed to avoidance behavior and one asks if the severity of the 'hole' has reduced over time as animals become habituated to the disturbance. To explore the possibility of temporal trends in the 'hole' we tested the significance of two interaction terms involving year: interactions with 'local hole proximity' and with 'big box'.

There was no indication that the effect of 'local hole proximity' varied with time. However, adding an interaction term between 'big box' and year yielded a large reduction in AIC compared to the simpler model with no interaction. Interpreting this interaction effect is tricky. Among the possible scenarios about temporal changes in the 'big box' effect are:

1. **Habituation:** There has been below-average relative density in the 'big box' over the period but this scarcity weakened over time.

2. **Intensification:** There has been below-average relative density in the 'big box' over the period and this scarcity intensified over time.

3. **New Effect:** There was above-average relative density in the region in early years but relative density has decreased over time to the extent that below-average relative density has existed in recent years.

4. **Other:** There are other scenarios about temporal variation that are not pertinent to this discussion due to the absence of any support for them in our findings below.

The bootstrap yielded 74.5% of simulations supporting the intensification scenario (i.e., a negative interaction parameter) and 25.5% supporting the habituation scenario. This provides no evidence of any consistent time trend in whale density in the 'big box' region.

However, the habituation hypothesis can be split into two possibilities: 1a) Habituation to the point where there is average or above-average relative abundance in recent years, and 1b) Habituation with the relative abundance in recent years remaining below average. Among 1a, 1b, 2 and 3, only 1a yields a conclusion that unusually low local abundance in the 'big box' no longer persists. Hypothesis 1a is rejected by the bootstrap ($p<0.001$). Additionally, 83% of bootstrap simulations supported 1b or 2, with the rest supporting scenario 3. Since 17% of scenarios had above-average relative abundance in the 'big box' in early years yet finished with below-average relative abundance in recent years, we can identify the year in which the transition occurred. We found that 95% of simulations showed the transition from over- to under-abundance were in 1986 or earlier. The latest apparent transition was 1991.

In summary, the relative abundance in the 'big box' is lower than outside it, and this phenomenon has persisted over time after correcting for the other factors in our model. There is no significant evidence that the unusual scarcity in the 'big box' is increasing or decreasing over time, but there is evidence that the phenomenon has been present or at least emerged in the late 1980's.

The previous 'hole contrast' finding supporting 'piling up' to the east and scarcity to the west could suggest reaction to an acute avoidance. However the above findings about the apparent lack of interaction between 'big box' and year—and the early onset of 'big box' scarcity—do not support a scenario of recent avoidance behavior. Instead, they suggest that scarcity in the 'big box' could be the manifestation of some persistent factor that predates many recent acute disturbances like those that could be attributable to industrial activity. We are unable to reconcile these two findings or to discard either hypothesis.

3.3 Sensitivity analysis

The BWASP dataset is not ideal. As noted in the introduction, survey effort was not located purely independently of animal abundances. Regardless of animal abundances, flights did not cover the entire survey region evenly: effort was lower in regions near the 30-minute section edges due to the transect design. Moreover, the transects should ideally be perpendicular to the general (east-west) axis of migration. This last concern is particularly relevant for bowhead whales because queues of migrating whales are not uncommon and could complicate relative abundance estimates when a transect traveled along a queue. Such concerns about the sampling scheme motivated sensitivity analysis. We analyzed several artificial datasets to assess sensitivity, and we report three main findings here.

In the first artificial dataset ('Western Oversampling') we identified all waits west of the median location. The 'Western Oversampling' dataset consisted of the original data supplemented by two more copies of each western record. This scenario was designed to mimic the situation where surveyors applied preferential effort in some geographical regions for reasons unrelated to whale abundance. Aside from obvious cases, note that edge effects from survey design are an example of such a scenario.

In the second artificial dataset ('Short Wait Oversampling') we identified all waits exceeding the median. The 'Short Wait Oversampling' dataset consisted of the original data

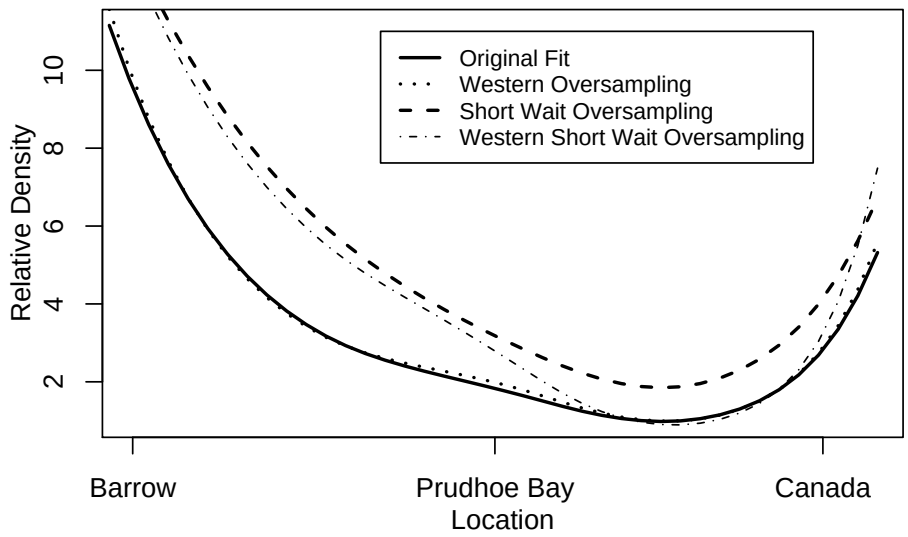

Figure 6: Comparison of profiles of estimated relative densities for three sensitivity analyses. The solid line is the fit to the original dataset. See the text for description of the three other fits.

supplemented by two more copies of each record with a waiting time below the median. Recall that short waits correspond to higher animal density. This scenario was designed to mimic the situation where effort is intentionally concentrated where the whales are believed to be present (more short waits); it is also conceptually analogous to targeting areas of believed scarcity (more long waits).

The third artificial dataset ('Western Short Wait Oversampling') consisted of the original dataset and two additional copies of any record triplicated in *both* of the previous datasets. In other words, short waits in the west were triplicated.

Figure 6 compares fits for relative density as a function of location along the shore, profiled at 30m depth for September 25, 2003. The dotted, dashed, and light dashed lines show the fits to the 'Western Oversampling', 'Short Wait Oversampling', and 'Western Short Wait Oversampling' datasets, respectively. The solid line is the fit to the original dataset. Figure 6 shows that the 'Western Oversampling' fit is nearly superimposed over the original fit. The 'Short Wait Oversampling' fit is similar in shape but shows consistently higher relative densities everywhere. The 'Western Short Wait Oversampling' fit is qualitatively similar to the others, but the ratio of the relative density in the west compared to the east is increased compared to the original fit.

The 'Western Oversampling' result can be understood using the analog of simple linear regression. If all the points on the left half of the regression scatterplot are triplicated, the fitted line should remain approximately unchanged. The 'Western Oversampling' sensitivity result in Figure 6 confirms that flying disproportionately in certain geographic regions should not, by itself, substantially alter relative abundance estimates. The nature of our analysis is that its estimation relies on information akin to sightings per unit effort, not raw counts of sightings.

The 'Short Wait Oversampling' results are also easily understood. In the BWASP dataset, short waits are distributed roughly uniformly across longitudes. Therefore, the excess short waits yield higher estimates of relative abundance without causing much change to the shape of the fit. Since relative abundance is determined up to a multiplicative constant, the result seen in Figure 6 would have no real impact on conclusions.

In the third case ('Western Short Wait Oversampling'), the artificial dataset has more short waits overall but proportionally fewer short waits in the eastern half of the region compared to the ratio in the original dataset. Figure 6 shows increased density isolated in the west. This result confirms that when whales are particularly abundant in a specific area and survey effort is particularly focused on that area, bias is possible.

Overall, these sensitivity results are fairly reassuring for creating maps like Figures 4 and 5. Not only to they suggest that our estimation method responds in the expected way to various sorts of changes to the data, but also they partially allay some concerns about the BWASP analysis. Non-uniform spatial coverage is only a limited concern except when effort is related to systematic patterns of animal presence. Even then, in our example, the qualitative shape of the estimated map is not greatly affected despite *tripling* the signal for high abundance in a target region.

4 Discussion

4.1 Model structure

The model of piecewise homogeneity merits further consideration. This approximation to full inhomogeneity is not tied to an explicit sampling resolution, in that the locations of covariate observations need not be linked to the scale over which they are likely to influence abundance. In fact, BWASP covariates were generally recorded substantially more frequently than sightings. This may partially mitigate the concern because it allows (through local averaging) some influence of smaller-scale covariate variation without precluding large-scale effects.

With respect to the assumption of piecewise homogeneity, animal density might reasonably be considered approximately constant over short waits, however for sightings very far apart the assumption of homogeneity is less credible. The same issue arises for covariates and sighting conditions: these are unlikely to be constant over long inter-sighting distances. HB propose an iterative method for adjusting for this concerns. Their approach relies on a nonlinear transformation of waiting distances to adjust for the estimated densities at the sighting locations. We have not implemented any adjustment for potential inhomogeneity in density between sightings here.

We have implemented an alternative approach (described above) for addressing inhomogeneity in covariates. For variables used to explain waiting periods, we calculated the covariates as weighted averages over the region preceding the sighting. This is consistent with our model which assumes piecewise homogeneity in presence. In contrast, variables related to sightability are used in a separate analysis here (and separately by Givens et al.

(2009)) to estimate a detection function for sightings. For these covariates, averaging on both sides of a sighting is sensible because a nearby region surrounding the sighting provides better information about sighting conditions at the point of detection than does a region containing points distantly preceding the sighting. Covariate averaging for sightability covariates is not required. However it takes advantage of much more available information, dilutes the effect of data errors, and allows substitution for missing values. If averaging of extra available covariates is to be used, our approach is more consistent with the standard situation when the covariates are only observed at the sighting location. Specifically, the detection function expresses the conditional probability of sighting an object given that the object is at a certain distance from the plane and given the covariate values at that sighting. The covariates that pertain are those that modify this relationship between distance and detection.

This discussion of waiting period homogeneity and sighting-specific density estimates raises a second concern. Consider a sighting following a very long preceding wait which is then followed by several subsequent sightings with very short waits. This would occur when there is a step function in relative density. Under the homogeneity assumption, the very long wait yields a low density estimate at the point of the step, rather than, say, at the midpoint of the long wait. Since the step occurs at a sighting, it is reasonable to view the sighting as originating from the beginning of the high density region. Under this view, the density estimate at such a step is biased downward because it is based completely on the preceding wait and ignores the upcoming sightings. In this scenario the bias is real but it should be put in context. Consider a bivariate scatterplot moving average smoothing problem. The nature of the smoother is that it will be biased downward near local maxima and biased upward near local minima (e.g., Givens and Hoeting, 2005). The same sort of effect will be seen with a step function: the smoothed estimate will blur the transition, with upward bias before the upward step and downward bias after it. Thus the concern about estimating animal density from waits in a region of abrupt density change is nothing more than an example of the same issue statisticians confront in many types of model-fitting. There is a trade-off between bias and variance, with some bias being tolerated in order to fit a smooth model to facilitate understanding of the underlying process that generated the data.

4.2 Related methods

A wide variety of other well-known methods are available for the analysis of spatial data. For example, the approach of Baddeley and Turner (2005) is similar to our method to the extent that it estimates the intensity parameter of an inhomogeneous Poisson process as a log linear function of covariates. In that implementation, however, the covariates must have a fixed spatial reference. The BWASP data include variables like year, however. The year may, in principle, be observed multiple times at the same spatial point with differing values. This is not compatible with their approach. One workaround would be to fit separate models for each year. This would reduce sample sizes by about 96% on average and eliminate any smooth interannual trends. Furthermore, the day variable presents the same problem, as do

variables relating to effort such as waiting distance, sightability, or whatever similar term might be developed. On the other hand, the Baddeley and Turner (2005) approach has a natural way to model clustering. This is appealing despite the fact that cluster size was found to be unrelated to sighting probability and cluster presence in our analysis. Their approach also enables the modeling of spatial stochastic dependence between points.

4.3 Conclusion

In conclusion, the BWASP data have presented a context for reconsideration of the spatial line transect model of Hedley and Buckland (2004). Several necessary and useful revisions and additions were identified for addressing issues like clustered observations and censored waits which were rampant in the BWASP data. Accordingly, a generalization of the conceptual model and an adaptation of a GAM estimation strategy were developed. Two important concerns about this approach are its sensitivity to deviations from a purely random effort distribution and its reliance–in the simplified GAM approach–on the assumption of inter-sighting piecewise homogeneity. Investigations showed that sensitivity to sampling biases can be comparatively low in the middle regions of the data. Concern about piecewise homogeneity is reasonable but can be viewed in the context of a bias-variance trade-off. For the BWASP data, the approach yielded very informative estimates of wide-scale spatio-temporal patterns in bowhead migration. These estimates were consistent with scientific conventional wisdom and hunters' traditional knowledge, while providing new spatio-temporal detail. We also illustrated how localized abundance variation could be investigated after controlling for the large-scale migration patterns. As examples of smaller-scale pattern detection, we found significant evidence of localized scarcity and localized 'piling up'. Possible reasons for such findings include anthropogenic impacts, variation in environmental conditions, whale behavior, or other unknown factors.

REFERENCES

Angliss, R. P., Rugh, D. J., Withrow, D. E., and Hobbs, R. C. (1995). Evaluations of photogrammetric length measurements of the western Arctic stock of bowhead whales (*balaena mysticetus*). *Report of the International Whaling Commission*, 45:313–24.

Baddeley, A. and Turner, R. (2005). Spatstat: an R package for analyzing spatial point patterns. *J. Stat. Software*, 12:1–42.

Buckland, S. T., Anderson, D. R., Burnham, K. P., Laake, J. L., Borchers, D. L., and Thomas, L. (2001). *Introduction to Distance Sampling: Estimating Abundance of Biological Populations*. Oxford University Press, Oxford, UK.

Buckland, S. T., Anderson, D. R., Burnham, K. P., Laake, J. L., Borchers, D. L., and Thomas, L. (2004). *Advanced Distance Sampling: Estimating Abundance of Biological Populations*. Oxford University Press, Oxford, UK.

Burnham, K. P. and Anderson, D. R. (1998). *Model Selection and Inference: a Practical Information-Theoretic Approach*. Springer-Verlag, New York, NY.

George, J. C., Zeh, J., Suydam, R., and Clark, C. (1994). Abundance and population trend (1978-2001) of the western Arctic bowhead whales surveyed near Barrow, Alaska. *Marine Mammal Science*, 20:755–773.

Givens, G. H. and Hoeting, J. A. (2005). *Computational Statistics*. Wiley, New York.

Givens, G. H., Hoeting, J. A., and Beri, L. (2009). Factors that influence aerial line transect detection of Bering-Chukchi-Beaufort Seas bowhead whales. *J. of Cetacean Research and Management*. submitted.

Griffiths, W. B. and Buchanan, R. A. (1982). Characteristics of bowhead feeding areas. Pages 347-455 in W.J. Richardson, ed., Behavior disturbance responses, feeding of bowhead whales, *Balaena mysticetus*, in the eastern Beaufort Sea, 1980-81. Report by LGL Ecological Research Associates for US Bureau of Land Management. NTIS No. PB86-152170.

Hedley, S. L., Buckland, S. T., and Borchers, D. L. (2004). *Spatial distance sampling models*, pages 48–70. Advanced distance sampling: estimating abundance of biological populations. Oxford University Press, Oxford UK. S. T. Buckland, D. R. Anderson, K. P. Burnham, J. L. Laake, D. L. Borchers, and L. Thomas (Eds.).

Manly, B. F. J., Moulton, V. D., Elliot, R. E., Miller, G. W., and Richardson, W. J. (2007). Analysis of covariance of fall migrations of bowhead whales in relation to human activities and environmental factors, Alaskan Beaufort Sea: Phase I, 1996-1998. OCS Study 2005-033; LGL Rep. TA2799-3. Rep. from LGL Ltd., King City, Ontario, and WEST Inc., Cheyenne, Wyoming, for US Minerals Management Service, Anchorage Alaska. 128pp.

McCullagh, P. and Nelder, J. A. (1989). *Generalized Linear Models*. Chapman & Hall, New York.

Monnett, C. and Treacy, S. D. (2005). Aerial surveys of endangered whales in the Beaufort Sea, fall 2002-2004. OCS Study MMS 2005-037, Anchorage, Alaska, 153pp.

Moore, S. E. and Reeves, R. R. (1993). *Distribution and Movement*, chapter 9, pages 313–386. Special Publication No. 2 of the Society for Marine Mammalogy. J.J. Burns, J.J. Montague and C.J. Cowles (Eds).

Richardson, W. J. and (Eds), M. T. W. (2004). Monitoring of industrial sounds, seals, and bowhead whales near BP's Northstar oil development, Alaskan Beaufort Sea, 1999-2003. LGL Report TA4002. Rep. from LGL Ltd. (King City, Ontario), Greeneridge Sciences Inc. (Santa Barbara, CA) and WEST Inc. (Cheyenne, WY), for BP Explor. (Alaska) Inc. Anchorage, AK.

Rugh, D., DeMaster, D., Rooney, A., Breiwick, J., Shelden, K., and Moore, S. (2003). A review of bowhead whale (*balaena mysticetus*) stock identity. *Journal of Cetacean Research and Management*, 5:267–279.

Satellite, N. and Information Service, N. G. D. C. (2008). Bathymetry. http://www.ngdc.noaa.gov/mgg/bathymetry/relief.html.

Schick, R. S. and Urban, D. L. (2000). Spatial components of bowhead whale (Balaena mysticetus) distribution in the Alaskan Beaufort. *Sea. Can. J. Fish. Aq. Sci.*, 57:2193–

2200.

Thomas, L., Laake, J. L., Strindberg, S., Marques, F. F. C., Buckland, S. T., Borchers, D. L., Anderson, D. R., Burnham, K. P., Hedley, S. L., Pollard, J. H., Bishop, J. R. B., and Marques, T. A. (2006). Distance 5.0. release 2. Research Unit for Wildlife Population Assessment, University of St. Andrews, UK. http://www.ruwpa.st-and.ac.uk/distance/.

Treacy, S. D. (2002). Aerial surveys of endangered whales in the Beaufort Sea, fall 2001. OCS Study MMS 88-0030, Anchorage, Alaska.

Treacy, S. D., Gleason, J. S., and Cowles, C. J. (2006). Offshore distances of bowhead whales (*balaena mysticetus*) observed during fall in the Beaufort Sea, 1982-2000: an alternative interpretation. *Arctic*, 59:83–90.

Wartzok, D., Watkins, W. A., Würsig, B., Guerrero, J., and Schoenherr, J. (1990). Movements and behaviors of bowhead whales. Report to Amoco Production Company, P.O. box 800, Denver, CO 80201. 197pp.

Wood, S. N. (2000). Modeling and smoothing parameter estimation with multiple quadratic penalties. *Journal of the Royal Statistical Society, Series B*, 62:413–428.

Wood, S. N. (2004). Stable and efficient multiple smoothing parameter estimation for generalized additive models. *Journal of the American Statistical Association*, 99:673–686.

Wood, S. N. (2007). The mgcv package. Reference manual available from http://www.maths.bath.ac.uk/∼sw283/simon/mgcv.html.

Würsig, B. and Clark, C. (1993). *Behavior*, pages 157–199. The Bowhead Whale. Lawrence, KS. J.J. Burns, J.J. Montague and C.J. Cowles(Eds.), Special Publication Number 2, Society for Marine Mammalogy.

Zeh, J. E., Clark, C. W., George, J. C., Withrow, D., Carroll, G. M., , and Koski, W. R. (1993). *Current population size and dynamics*, pages 409–489. The Bowhead Whale. Society for Marine Mammaology. J.J. Burns, J.J. Montague and C.J. Cowles(Eds.), Special Publication Number 2.