Environmetrics 00, 2–31

DOI: 10.1002/env.XXXX

Horvitz-Thompson Whale Abundance Estimation Adjusting for Uncertain Recapture, Temporal Availability Variation and Intermittent Effort

Geof H. Givens[a*], Stacy L. Edmondson[b], J. Craig George[c], Robert Suydam[c] Russell A. Charif[d], Ashakur Rahaman[d], Dean Hawthorne[d], Barbara Tudor[c], Robert A. DeLong[e] and Christopher W. Clark[d]

Summary: A Horvitz-Thompson type estimator is introduced to estimate total abundance of the Bering-Chukchi-Beaufort Seas population of bowhead whales using combined visual and acoustic location data. The estimator divides sightings counts by three correction factors that are themselves estimated from various portions of the data. The first correction models how detection probabilities depend on covariates like offshore distance and visibility. The second correction adjusts for availability using the acoustic location data to estimate a time-varying smooth function of the probability that animals pass within visual range of the observation stations. The third correction accounts for whales passing during periods when one or both sighting stations were temporarily closed down. We derive an asymptotically unbiased estimator of abundance incorporating all these components, and a corresponding variance estimate. Correcting the count of 4,011 observed whales yields a 2011 abundance estimate of 16,820 with a 95% confidence interval of (15,176, 18,643) and an estimated annual rate of population increase of 3.7% (2.9%, 4.6%). These results are indicative of very low conservation risk for this population under the current low levels of aboriginal hunting permitted by the International Whaling Commission. Although few other capture-recapture surveys will confront exactly the same set of challenges addressed here, many studies face one or more issues that could be resolved by adapting portions of our approach or relevant underlying concepts thereof. Moreover, the generic estimator we derive represents an improved way to handle random correction factors rather than assuming fixed values.

Keywords: capture-recapture uncertainty, misidentification, whale abundance

This paper has been submitted for consideration for publication in *Environmetrics*

1. INTRODUCTION

Bowhead whales (*Balaena mysticetus*) are a large baleen whale species that live in arctic waters and include a large, well-studied population in the Bering, Chukchi and Beaufort Seas. Native Alaskans conduct limited subsistence hunting of this migrating population from several remote coastal villages, with harvest levels determined by the International Whaling Commission.

In the spring of 2011, a major multi-faceted program of research on this whale population was undertaken, including ice-based visual counting, underwater acoustic monitoring, aerial photo-identification, satellite tagging and biopsy sampling. In this paper, we use the visual and acoustic data to estimate total bowhead population abundance and update the estimate of population increase rate. Although the dataset arises from multiple visual and acoustic detection opportunities, estimation is not straightforward because the survey scheme violates several precepts of standard capture-recapture analysis. Specifically, (i) the identification of recaptures is prone to error, (ii) there is smooth temporal variation in the availability of whales to be detected within the visual detection range, and (iii) weather or other factors sometimes compel one or both sighting stations to temporarily cease operations while whales migrate past at a time-varying rate. Also, detection probability must be estimated from a set of covariates.

Our presentation is organized as follows. In the next section, we describe the survey design and data. Detailed expositions of the survey protocols and available datasets are given by George et al. (2013) and Clark et al. (2013). Our statistical methods are developed

[a] *Givens Statistical Solutions LLC, 4913 Hinsdale Dr., Fort Collins CO 80526. E-mail: geof@geofgivens.com*
[b] *Mathematics Department, Whitman College, 345 Boyer Avenue, Walla Walla, WA 99362*
[c] *North Slope Borough, Department of Wildlife Management, Barrow AK 99723, U.S.A.*
[d] *Bioacoustics Research Program, Cornell Laboratory of Ornithology, Cornell University, Ithaca, NY 14850 U.S.A.*
[e] *DeLongView Enterprises, Box 85044, Fairbanks, AK 99708, U.S.A.*
[*] *Correspondence to: Geof H. Givens, Givens Statistical Solutions LLC, 4913 Hinsdale Dr., Fort Collins CO 80526. E-mail geof@geofgivens.com*

in Section 3. Analysis results follow in Section 4. The final section of the paper provides discussion and context for our findings.

2. DATASETS

Our analyses use two datasets collected in the spring of 2011. A visual sightings dataset contains spatio-temporal data about whale sightings made from two observer stations at fixed locations, along with various covariates (e.g., visibility) and certain other data. An acoustic dataset was derived from whale sounds as recorded on an underwater array of 3-6 acoustic recording devices. A near-field beam forming spatial energy maximization approach was used to estimate the spatial location of each sound (Clark et al., 2013). These location estimates (and times) comprise our acoustic dataset. The acoustic detection region is larger than the visual detection region and encompasses it.

The 2011 visual and acoustic data collection season ran from April 4, when the first visual watch was conducted on the ice edge, until July 27 when acoustic recording ended. The first bowhead whale was seen on April 9. Our analyses are limited to a shorter season described below that includes the vast majority of visual sightings.

2.1. Visual Data

George et al. (2013) explain the details of the visual survey. Briefly, two visual observation perches were erected on a pressure ridge on the shore-fast ice near the water edge. The perches were 39.4 meters apart, which was sufficiently distant that observers on one perch operated wholly independently from those on the other. The south perch was designated primary, and we attempted to staff it with rotating teams of at least 3 observers at all times, except as limited by safety concerns and weather. The north perch was staffed intermittently for periods of 'independent observer' or 'IO' effort. The online supplementary material for this article provides more details about IO timing.

The visual data were collected by ice-based observers sighting whales as they migrated northeast along the shore-fast ice edge past Barrow, Alaska. Observers saw 3379 'New' and 632 'Conditional' whales from the primary observation perch. George et al. (2013) explain the distinction between New and Conditional sightings. Essentially, when observers are unsure whether a whale has been previously sighted, it is labeled as Conditional. The implications of this distinction are discussed later.

For the purpose of analysis, the 2011 visual census is defined to have begun at 14:35 local time on April 13, 2011, and ended at 16:00 on June 1, 2011. These are, respectively, the beginning of the first watch session (from the primary perch) and the end of the last watch session during which a whale was seen. After June 1, it was too dangerous to continue visual effort. Many of our plots display data by hours of the year; in these units the season ran from 2462.583 to 3640.

The visual survey data have been used to estimate the probability of detecting a whale or group given that it is present (Givens et al., 2014). They also provide the counts that are the foundation of our total abundance estimate.

2.2. Acoustic Data

The acoustic data are used to estimate the proportion of whales that migrate within visual range. This analysis provides an important correction factor for the total abundance estimate.

The acoustic dataset was derived from continuous sound recordings from an array of up to six underwater acoustic recorders that were deployed near the ice edge in the vicinity of the visual observation perches and recovered later that summer. Clark et al. (2013) describe the details. From these recordings, a subsample of time periods was examined to identify whale calls and song. The raw data from the recorder array were used by Clark et al. (2013) to estimate spatial locations and corresponding 95% confidence regions. Hereafter, we take their results at face value and refer to these processed data as the 'acoustic location' estimates. A total of 22,426 bowhead vocalizations yielding acoustic location estimates were collected (of which only a relevant portion were used for analysis as discussed later). There is no way to know how many whales are represented by this large number of vocalizations since during their passage through the acoustic monitoring area some whales will vocalize more frequently than others and some may not produce a single sound. Also, it is extremely difficult to pinpoint which sounds are associated with a specific visual sighting.

Figure 1 sketches the survey layout. Although this figure is only roughly scaled and oriented, true north is toward the top and the ice edge is represented by a line that runs from southwest to northeast. Migration proceeds roughly parallel to the ice edge. The two perches are shown as small squares, and the six acoustic recorders are stars.

[Figure 1 about here.]

The larger semicircle in Figure 1 is 20 km from the array centroid. When an acoustic location was estimated to be more than 20 km offshore, the offshore distance was set equal to 20 km. This was done because the range estimator was considered to provide an imprecise (and large) distance for such cases, even though the bearing estimate would be reliable. The array axis is defined by the line between the southwestern-most and northeastern-most recorders. The region within 30 degrees of the array axis and beyond the ends of the array is called the endfire zone. Distance estimates for locations in the endfire zone are considered unreliable due to the geometry involved, and those data are discarded.

The north-easternmost and southwestern-most recorders also determine the aperture of the acoustic array. Roughly, the array aperture is defined to be the length of the segment of the array axis between the ends of the array. The two parallel dotted lines that extend the aperture outward, perpendicular to the ice edge, define a strip called the aperture zone. Data within the aperture zone play an important role in the analysis below.

The smaller semicircle in Figure 1 is 4 km from the perches. This represents the practical limit of visual range, and only the sightings within this range (96%) are analyzed to estimate detection probability (Givens et al., 2014) and abundance (here). Accounting for rare sightings beyond the practical visual range is done via availability estimation discussed later.

2.3. Combined dataset

[Figure 2 about here.]

Figure 2 summarizes the visual and acoustic data used in our analyses. The horizontal dimension of this figure is time, which is indexed by hour on the bottom axis and calendar date on the top axis. The dual axes are for convenience: the two axes match and either may be used everywhere in the figure. The top portion of the plot shows the acoustic data and the vertical axis is distance from the perch. This shows only the data within the acoustic array aperture zone that were not excluded for data quality reasons. Each point corresponds to one acoustic location at a particular time and a particular distance from the ice edge. The shaded (blue) vertical stripes are times when the recordings were analyzed to estimate locations. About 28% of the analyzed season was examined. The lower portion of the plot shows the visual data. Counts of sighted whales are summarized by a (upside-down) histogram with black bars. The histogram bins are 6 hours wide. The shaded (red) vertical stripes correspond to periods with qualifying watch effort from the primary perch. About 45% of the analyzed season was covered with qualifying primary perch effort (see Section 3.1). Only sightings made from the primary perch during these times are counted in the abundance estimate. When the histogram bin edges extend outside the shaded stripes, it should be understood that all the sightings within the bin occurred within the stripe.

3. METHODS

In the following subsections we describe estimation of key quantities used in our abundance estimate. These components of our analysis are estimated using a variety of techniques including familiar models with new twists and novel approaches that are specialized for unusual aspects of the whale survey. Section 3.5 then presents the overall modeling framework with new estimators of abundance and variance, and their properties. Our estimator is applicable to the important and relatively common situation when (estimated) correction factors are subject to sampling variability and should not be considered constants.

3.1. Overview

Visual sightings data refer to groups of whales, although 83% of these groups were size 1. Although group memberships may vary during passage, the groups are conceived as being defined when they pass the perches. Let c_i represent a sighting group size, for $i = 1, \ldots, g$, where g is the number of groups sighted.

Whales are very difficult to see beyond 4 km, although some sightings can be made under the very best possible visibility conditions. Our analysis assumes that bowheads are only available to be seen by observers when they swim and surface within the 4 km radius visual detection zone. More distant sightings are truncated from the analysis. Let a_i denote the probability that the ith group was available. If the group is available for visual detection, it may or may not actually be seen from the primary perch. Define the detection probability p_i to be the conditional probability that the ith group was seen given that it was available. Let $\widehat{a}_i$ and $\widehat{p}_i$ denote estimators for a_i and p_i.

During some portions of the season, there was no observer effort because the perch was not staffed, visibility was poor or unacceptable, environmental conditions were unsafe, or wind had moved the sea ice so that it completely covered the survey region. In good conditions, there is usually an 'open lead'–a channel of open water between the shore-fast ice and floating ice–or nearly wide open water. Let $H_s = 1177.417$ denote the total number of hours during the season (i.e., from hour 2462.583 to 3640), and let H_w denote the total number of those hours for which observer watch effort was maintained during qualifying conditions. Since $H_w < H_s$, the abundance estimator must correct for periods of missed survey effort.

Denote the unknown total population size as N. Our abundance estimator employs a scaled modified Horvitz-Thompson approach (Borchers et al., 2002; Horvitz and Thompson, 1952). The abundance estimate is

$$\widehat{N} = \frac{1}{\widehat{E}} \sum_{i=1}^{g} \frac{c_i}{\widehat{a}_i \widehat{p}_i} = \widetilde{N}/\widehat{E}. \tag{1}$$

where $1/\widehat{E}$ is a correction for whales passing at missed times. For brevity, we will often refer to this as an effort correction, and it must be estimated because despite knowing the times when the perches were and weren't operational, the passage rate and number of whales passing during those times are unknown. The group sizes c_i used in (1) are only the sightings from the primary perch. The data from the second perch are used to estimate detection probability and the effort correction. The merit of this choice is discussed later.

In the supplementary material we derive the abundance estimator and its theoretical mean and variance as extensions to the results of Steinhorst and Samuel (1989). Our approach to variance estimation extends that of Wong (1996); see also Fieberg (2012). We also provide an asymptotically unbiased variance estimator to replace the biased estimator of Steinhorst and Samuel (1989). See Section 3.5 and the supplementary material for further details.

3.2. Availability estimation

We use the acoustic location data to estimate the a_i. The raw acoustic data are filtered to exclude the locations whose 95% confidence intervals for bearing extend greater than 22.5 degrees from the corresponding point estimate, locations in the 'endfire' regions, rare locations falling on the grounded ice or land, and locations identified during additional pre-processing by Clark et al. (2013) as almost certainly being additional sounds from the same whale. Here we examine only locations within the array aperture zone, at any distance from the ice edge (see Figure 1). Only these data are displayed in Figures 2 and 4.

Our use of the acoustic data and the aperture zone relies on several assumptions. We assume that on average, the number of locations at any distance is proportional to the number of whales passing at that distance (George et al., 2004, p. 762). Note that this does not imply that each whale is represented by only a single sound in the dataset. It follows that acoustic behavior does not systematically vary with distance offshore or vary between whales in any way that would bias estimation of the a_i. There are empirical data supporting this assumption. For example, an analysis of 'call tracks', i.e., a sequence of sounds whose characteristics enable it to be matched to a single, identifiable whale, indicates that the number of calls per track was essentially identical for distances less than 4 km and greater than 4 km. Finally, we assume that reduced acoustical detectability with increasing distance within the 20 km range analyzed here is ignorable. Detectability is related to array length and the wavelength of the sound. Using a commonly accepted rule of thumb, the effective range was over 500 km and the distortion at 4 km is negligible.

The acoustic data include estimated offshore distances $\widehat{d}_i$ (meters) and times t_i for $i = 1, \ldots, L$ locations. Each point is assigned a binary outcome $\widehat{b}_i$ that equals 1 if $\widehat{d}_i \leq 4000$ and 0 otherwise. It is important to understand that there is uncertainty in the $\widehat{d}_i$. Clark et al. (2013) describe how a two-dimensional confidence region is estimated for each location, and how this is converted to a confidence interval for each d_i. Although it might seem at first that sampling errors for the $\widehat{d}_i$ would be skewed to the left, this is, to the first order, not true. The correlation sum estimator those authors use finds a single energy maximum in space and is not based on sound arrival times at sensors within the array. It is therefore reasonable to proceed with the assumption that the offshore distance error distribution is symmetric. The supplementary material discusses several other assumptions.

We use these results to calculate a weight w_i for each $\widehat{b}_i$. Specifically, we convert the Clark et al. (2013) results to approximate confidence intervals for distances by assigning to each offshore distance a normal distribution, centered at $\widehat{d}_i$ and having a standard error implied by those results. We then define $w_i = \left|P[\widehat{d}_i < 4000] - 0.5\right|$. Thus the weights are intended to be proportional to the probability that $\widehat{b}_i$ is correct considering the inherent variability in the location estimates.

Recall that our goal is to estimate the proportion of whales that are available to be visually detected within visual range. To be available, the whale must surface at least once within the 4 km semicircle in Figure 1. Conceptually, we estimate this by examining the proportion of acoustic locations inside the aperture zone that are within 4 km of the ice edge. The boundaries of the aperture zone are designated in Figure 1 by the two long dotted parallel lines passing through the array ends and perpendicular to the ice edge. These lines define a strip, and the innermost 4 km of this strip defines a rectangular box where whales may swim through the visual detection zone. Graphically, our estimate compares the number of acoustic locations in this box to the number in the entire strip. In concept, this comparison is the same one used by George et al. (2004).

A whale just less than 4 km from the visual perch has some nonzero probability of passing through the aperture zone yet never surfacing in the visual detection zone, since the nearest 4 km of the aperture zone is a box but the visual detection zone is semicircular. In fact, every whale has some chance of doing this if it holds its breath long enough. A Monte Carlo experiment described in the supplementary material shows that these issues should have negligible impact on the results. Our approach also maintains consistency with analyses of past surveys.

We adopt a weighted quasi-binomial generalized additive model (GAM) for the b_i data (Wood, 2004, 2006, 2011). The model was fit using the `mgcv` package in the **R** computing language (R Core Team, 2015). Defining $a_i = P[b_i = 1]$, we model

$$\log\left\{\frac{a_i}{1-a_i}\right\} = f_a(t_i) \tag{2}$$

where f_a is a penalized regression spline formed from a thin plate regression spline basis, which is the default in the `mgcv` package. The model fitting employed our weights, w_i. The number of knots was set at k=20, which allows good fidelity to the data at a temporal frequency and resolution consistent with observer opinions about the rate at which the offshore distribution of whales changes, without over-fitting. Also, in a plot of k versus the unbiased risk estimator criterion (not shown here), there is a clear, abrupt 'knee' at $k = 20$, which we interpret as an empirical indicator of a good choice. The default generalized cross-validation method was used to choose the smoothness penalty.

This model can be re-expressed in terms of the underlying spline basis functions. Let $\mathbf{Z}$ represent the (transposed) model matrix fashioned from the basis, with one row per basis function and the ith column $\mathbf{Z}_i$ corresponding to the ith case. Then we may write the model as

$$\log\left\{\frac{a_i}{1-a_i}\right\} = \mathbf{Z}_i^T\boldsymbol{\alpha} \tag{3}$$

where $\boldsymbol{\alpha}$ is a column vector of parameters. Fitting equation (2) amounts to estimating $\boldsymbol{\alpha}$. The asymptotic distribution of the parameter estimates $\widehat{\boldsymbol{\alpha}}$ can be summarized by $\widehat{\boldsymbol{\alpha}} \sim N(\boldsymbol{\alpha}, \boldsymbol{\Psi})$. Technically, this is a limiting Bayesian posterior distribution, but no prior information about $\boldsymbol{\alpha}$ or the a_i is incorporated in the analysis beyond the smoothness penalty; see Wood (2006). An estimated covariance matrix $\widehat{\boldsymbol{\Psi}}$ is obtained while fitting this GAM.

3.3. Detection probability estimation

Givens et al. (2014) describe estimation of the p_i. Their approach is complex, so we offer only a brief summary here.

Those authors applied a weighted Huggins (1989) model to capture-recapture data from the two-perch independent observer data. A critical component of their analysis was matching, i.e., the determination of whether a whale seen at one perch was the same individual as a sighting from the other perch. This process is described by George et al. (2012) and Givens et al. (2014), but is not relevant to the analyses in this paper beyond its contribution to detection probability estimation.

The estimation approach modeled the ith group as having a detection probability p_i. Then the conditional probability of sighting the group only at the primary perch is $p_i(1-p_i)/d_i$ where $d_i = 1-(1-p_i)^2$ is used because the model is conditioned on seeing the group at least once. The probability of sighting the group only at the second perch is the same, and the probability of sighting the group at both perches is p_i^2/d_i.

Many covariates were recorded along with each sighting. We can express these data in a (transposed) model matrix $\mathbf{X}$ with the ith column $\mathbf{X}_i$ corresponding to the ith sighting. After excluding data from the worst two visibility categories, the only covariates that significantly affected p_i were distance of the sighting from the perch, lead condition, and number of whales in the group. A generalized linear model was used to model the dependence:

$$\log\left\{\frac{p_i}{1-p_i}\right\} = \mathbf{X}_i^T\boldsymbol{\beta} \tag{4}$$

where $\boldsymbol{\beta}$ is a parameter column vector to be estimated. Estimated detection probability for a sighting, $\widehat{p}_i$, was derived from the parameter estimates:

$$\widehat{p}_i = \frac{\exp\{\mathbf{X}_i^T\widehat{\boldsymbol{\beta}}\}}{1+\exp\{\mathbf{X}_i^T\widehat{\boldsymbol{\beta}}\}}.$$

Givens et al. (2014) used a weighted likelihood estimation method, extending the basic model above to account for three uncertainties:

- Some sightings at one occasion may be unintentional resightings of a group already seen at the same occasion (called 'Conditional' whales; the rest are 'New').
- The identification of a recapture is uncertain and is given a confidence rating. When a recapture is falsely declared, the constituent data actually comprise two non-recaptured sightings.
- Sightings do not enjoy equal opportunity to be discovered as recaptures because periods when the second observer team was not operating produce only partial data for use in the matching process.

A weighted fit to the model yields parameter estimates $\widehat{\boldsymbol{\beta}}$ and the asymptotic result $\widehat{\boldsymbol{\beta}} \sim N(\boldsymbol{\beta}, \boldsymbol{\Phi})$. An estimate of the covariance matrix is obtained as part of weighted fitting of the detection probability model; denote this $\widehat{\boldsymbol{\Phi}}$.

3.4. Whales passing at missed times

Figure 2 shows the periods of visual effort during the season during qualifying visibility and lead conditions. To correct for periods without effort, it does not suffice to add up missed clock time—we must account also for the passage rate of whales during the missed periods.

To do this, we begin by recalling from equation (1) that $\widehat{N}$ involves a sum of terms

$$\widehat{h}_i = c_i/\widehat{a}_i\widehat{p}_i,$$

which we call Horvitz-Thompson contributions since the $\widehat{h}_i$ represent the estimated number of whales that the ith sighting contributes to the overall abundance estimate (uncorrected for effort). Figure 3 plots the Horvitz-Thompson contributions against time during the season.

Note that whale abundance is symbolized in this plot by *both* the density of points and the magnitudes of individual points.

[Figure 3 about here.]

Let $f_r(t)$ denote the passage rate of whales past the census area, so that the total number of whales passing the perch at any distance, detected or unseen, between time t_1 and t_2 is $\int_{t_1}^{t_2} f_r(t)dt$. Let S and W denote the sets of time periods corresponding to the total analyzed survey season and periods of qualifying watch effort, respectively. Then the proportion of the total population passing Barrow during the season that passed during periods of qualifying watch effort is

$$E = \int_W f_r(t)dt \bigg/ \int_S f_r(t)dt$$

and if we can estimate this quantity then the desired effort correction factor in equation (1) is $1/\widehat{E}$. This approach relies on the fact that passage rate is not correlated with observer presence, as shown from acoustic and aerial observations and the traditional knowledge of native hunters in the region. We also assume that the model of a smoothly varying passage rate over all periods of the day is reasonable.

To estimate f_r and hence E, we bin the $\widehat{h}_i$ into 12-hour time blocks, $\mathcal{B}_1, \ldots, \mathcal{B}_{101}$ and define $\widehat{H}_j$ to equal the sum of all $\widehat{h}_i$ that occurred during block $\mathcal{B}_j$. Thus, $\widehat{H}_j$ is the total Horvitz-Thompson contribution for the jth block, i.e., an estimate of the total number of whales passing during that block during times of qualifying effort in the ith block. Let T_j denote the amount of qualifying watch effort during the jth block and let the blocks be referenced by their temporal midpoints t_j. Then define $\widehat{R}_j = \widehat{H}_j/T_j$, to be the number of passing whales per qualifying watch hour in the jth block.

We adopt a shifted gamma generalized additive model with log link to model the block passage rates according to $\widehat{R}_j + s \sim \text{gamma}(\gamma_j, \eta_j)$ where the mean of $\widehat{R}_j + s$ is $\zeta_j = \gamma_j\eta_j$

and

$$\log \zeta_j = f_r^*(t_j)$$

where $f_r(t) = \exp\{f_r^*(t)\}$ is a smooth rate function. The shift constant s was necessary to cope with the fact that some $\widehat{R}_j = 0$; the value of s was chosen to maximize the explained deviance. The suitability of this model is discussed in the supplementary material.

We use the same GAM fitting tools and technical assumptions as previously used for modeling availability. Points are weighted proportionally to the T_j.

This model can be re-expressed in terms of a matrix $\mathbf{U}$ with one row per spline basis function and the jth column representing the jth block, using $\log \zeta_j = \mathbf{U}_j^T \boldsymbol{\gamma}$. The column vector of parameter estimates $\widehat{\boldsymbol{\gamma}}$ has a limiting posterior distribution $\widehat{\boldsymbol{\gamma}} \sim N(\boldsymbol{\gamma}, \boldsymbol{\Lambda})$ in the same sense as above. The covariance matrix estimate $\widehat{\boldsymbol{\Lambda}}$ is also derived during model fitting.

What remains is to estimate E. We set

$$\widehat{E} = \int_W \widehat{f}_r(t)dt \Big/ \int_S \widehat{f}_r(t)dt \tag{5}$$

where the integrals are approximated using Simpson's rule (e.g., Givens and Hoeting 2013). The subintervals in this numerical integration can be made sufficiently small so as to render the error in this approximation negligible.

Let $\widehat{\mathrm{var}}\{1/\widehat{E}\}$ denote the estimated variance of the correction factor estimator $1/\widehat{E}$. We estimate the variance using the parametric bootstrap approach recommended by Wood (2006, p. 202-3). Briefly, the GAM is first fit to the original data, then bootstrap iterations proceed as follows. Using the estimated mean function from the original fitted model, bootstrap response data are generated from the parametric (gamma) model. A new GAM is fit to these data to obtain a bootstrap estimate of the smoothing parameter. Next, a GAM is fit to the original data using the bootstrap smoothing parameter value. This produces one set of pseudo-estimates $\widehat{\boldsymbol{\gamma}}^*$ and $\widehat{\boldsymbol{\Lambda}}^*$. We performed 2500 bootstrap iterations.

Then, to simulate from the bootstrap distribution of $1/\widehat{E}$, we select at random one of the 2500 distributions $N(\widehat{\boldsymbol{\gamma}}^*, \widehat{\Lambda}^*)$ and sample a value $\boldsymbol{\gamma}^{**}$ from it. This value is used to obtain a bootstrap pseudo-value $\widehat{E}^*$ via equation (5). Finally, the sample variance of the values of $1/\widehat{E}^*$ is computed to produce $\widehat{\text{var}}\{1/\widehat{E}\}$.

Note that $1/\widehat{E}$ and its variance estimator are not statistically independent of the other key estimators ($\widehat{a}_i$, $\widehat{p}_i$ and $\widehat{\theta}_i$) in this paper. However, the nature of $\widehat{E}$ as an integral of a smooth function of the huge set of those quantities should provide reasonable justification to treat $\widehat{E}$ as approximately independent of our other estimators for our purposes.

3.5. Abundance estimation

Recall that the total abundance estimate can be written as $\widehat{N} = \widetilde{N}/\widehat{E}$, where $\widetilde{N}$ is the estimated total abundance of animals passing during times of observer (visual) effort and $1/\widehat{E}$ is the estimated correction factor accounting for whales passing at missed times. Then

$$\theta_i = \frac{1}{a_i p_i}$$

encapsulates availability and detectability correction factors so that

$$\widehat{N} = \frac{1}{\widehat{E}} \sum_{i=1}^{g} c_i \widehat{\theta}_i.$$

Let $\widehat{\theta}_i$ denote an estimator for θ_i.

In the supplemental material we derive asymptotically unbiased estimators for θ_i, $\widetilde{N}$ and for corresponding variances. Our approach employs the logit link relationships in our availability and detection probability models, the asymptotic normality of $\widehat{\boldsymbol{\beta}}$ and $\widehat{\boldsymbol{\alpha}}$, the bootstrap variance for $1/\widehat{E}$, properties of the lognormal distribution, and the approximation that $\widehat{\boldsymbol{\Phi}}$ and $\widehat{\boldsymbol{\Psi}}$ can be treated as known for large samples. Here we simply present the results.

For notational simplicity, it is useful to define some terms related to the linear predictors and covariance matrices in the generalized additive models for the visual and acoustic data. Specifically, define

$$
\begin{array}{lllll}
\mu_i = \mathbf{X}_i^T \boldsymbol{\beta} & \widehat{\mu}_i = \mathbf{X}_i^T \widehat{\boldsymbol{\beta}} & \phi_i = \mathbf{X}_i^T \widehat{\boldsymbol{\Phi}} \mathbf{X}_i & \phi_{ij} = \mathbf{X}_i^T \widehat{\boldsymbol{\Phi}} \mathbf{X}_j & \widetilde{\phi}_{ij} = \phi_i/2 + \phi_j/2 + \phi_{ij} \\
\eta_i = \mathbf{Z}_i^T \boldsymbol{\alpha} & \widehat{\eta}_i = \mathbf{Z}_i^T \widehat{\boldsymbol{\alpha}} & \psi_i = \mathbf{Z}_i^T \widehat{\boldsymbol{\Psi}} \mathbf{Z}_i & \psi_{ij} = \mathbf{Z}_i^T \widehat{\boldsymbol{\Psi}} \mathbf{Z}_j & \widetilde{\psi}_{ij} = \psi_i/2 + \psi_j/2 + \psi_{ij}
\end{array}
$$

using the notation established previously in this article. Note that terms such as ϕ_i and ψ_{ij} denote projections and quadratic forms related to the estimated covariance matrices, not individual terms therein. Since we treat the covariance matrices as known, we don't put hats on these expressions.

Then an asymptotically unbiased estimator of θ_i is

$$
\widehat{\theta}_i = (1 + \exp\{-\widehat{\mu}_i - \phi_i/2\}) \, (1 + \exp\{-\widehat{\eta}_i - \psi_i/2\}) \, .
$$

Further,

$$
\begin{aligned}
\widehat{\mathrm{var}}\{\widehat{\theta}_i\} &= \exp\{-2\widehat{\mu}_i - 2\phi_i\} \, (1 + 2\exp\{-2\widehat{\mu}_i - \widehat{\eta}_i - 2\phi_i - \psi_i\}) \, (\exp\{\phi_i\} - 1) + \\
&\quad \exp\{-2\widehat{\eta}_i - 2\psi_i\} \, (1 + 2\exp\{-\widehat{\mu}_i - 2\widehat{\eta}_i - \phi_i - 2\psi_i\}) \, (\exp\{\psi_i\} - 1) + \\
&\quad \exp\{-2\widehat{\mu}_i - 2\widehat{\eta}_i - 2\phi_i - 2\psi_i\} \, (\exp\{\phi_i + \psi_i\} - 1)
\end{aligned}
$$

is asymptotically unbiased for the variance of $\widehat{\theta}_i$. The supplemental material also gives an asymptotically unbiased estimator $\widehat{\mathrm{cov}}\{\widehat{\theta}_i, \widehat{\theta}_j\}$.

Using these results, we can derive the key asymptotically unbiased estimators: $\widetilde{N} = \sum_{i=1}^{g} c_i \widehat{\theta}_i$ and $\widehat{\mathrm{var}}\{\widetilde{N}\} = \widehat{V}_1 + \widehat{V}_2$ where $\widehat{V}_1 = \sum_{i=1}^{g} c_i^2 \left(\widehat{\theta}_i^2 - \widehat{\theta}_i - \widehat{\mathrm{var}}\{\widehat{\theta}_i\} \right)$. and $\widehat{V}_2 = \sum_{i=1}^{g} c_i^2 \widehat{\mathrm{var}}\{\widehat{\theta}_i\} + \sum \sum_{i \neq j}^{g} c_i c_j \widehat{\mathrm{cov}}\{\widehat{\theta}_i, \widehat{\theta}_j\}$. See the supplemental material for the details. Now $\widehat{N} = \widetilde{N}/\widehat{E}$ and we can estimate the variance of $\widehat{N}$ as the variance of the product of independent random variables:

$$\widehat{\text{var}}\{\widehat{N}\} = \frac{1}{\widehat{E}^2}\widehat{\text{var}}\{\widetilde{N}\} + \widetilde{N}^2\widehat{\text{var}}\{1/\widehat{E}\} + \widehat{\text{var}}\{\widetilde{N}\}\widehat{\text{var}}\{1/\widehat{E}\}. \tag{6}$$

For a simpler problem, Wong (1996) has demonstrated that it is better to estimate a confidence interval for N by applying a normal approximation to log abundance and then back-transforming the result. If we define $\widehat{CV}^2 = \widehat{\text{var}}\{\widehat{N}\}/\widehat{N}^2$, the estimated 95% confidence interval for N is $\left(\widehat{N}\exp\{-1.96\widehat{CV}\},\ \ \widehat{N}\exp\{1.96\widehat{CV}\}\right)$.

The counts c_i we use for this abundance estimate include both New sightings (whales definitely seen for the first time) and Conditional sightings (whales seen a second time from the same perch and observers are unsure whether the whale has been previously seen). Previous abundance estimates have always treated Conditional whales as half a whale each; we continue that tradition here.

We do not include whales seen only at perch 2. The reason for this is explained in Section 5.

3.6. Trend estimation

In this section we incorporate our abundance estimate into a longer time series of estimates in order to estimate population rate-of-increase, or trend. Heretofore trend has been estimated using a series of counts (scaled up to correct for detection probability) and availability estimates that are denoted N_4 and P_4, respectively, by Zeh and Punt (2005). The notation indicates that N_4 is the corrected count of whales sighted within 4km of the perch(es) and P_4 is the estimated proportion of whales that swim within that visual range. There are 11 years between 1978 and 2001 for which either N_4, P_4 or (usually) both have been obtained. This is a valuable time series from which we may estimate trend. Our approach is based on the method developed previously for this population (Cooke, 1996; Punt and Butterworth, 1999; George et al., 2004; Zeh and Punt, 2005).

The surveys between 1978 and 2001 are correlated because they share information about availability: the P_4 values for certain years were used to make abundance estimates for other years when no separate estimate of P_4 is available. The trend estimation approach we describe here accounts for the resulting correlation. It is a two-step procedure.

The first step is to estimate indices of abundance for all years when N_4 estimates are available (regardless of whether a corresponding P_4 is available). This estimation proceeds by fitting a model having three components. First, each observed log abundance is assumed to equal the sum of the true total log abundance in that year, the log proportion of the population within visual range in that year, and an independent normal error. Second, each observed log proportion within visible range is assumed to equal the sum of the corresponding true log proportion within visible range for that year and an independent normal error. Third, the true log proportion within visible range is assumed to equal a grand mean log proportion plus normal error. The second and third components introduce inter-annual process error. The overall model combining these three components is fit by restricted maximum likelihood. These abundances are indices created to 'share information' about P_4 for years in which no P_4 was directly estimated.

The second step of the process is to estimate trend using the fitted abundance indices. The trend can be estimated by fitting an exponential growth model using generalized least squares, incorporating the variance-covariance matrix of log abundances estimated in step 1 as the weighting matrix. A confidence interval for the trend estimate is calculated using asymptotic results.

Incorporating our new 2011 estimate into this procedure is not entirely straightforward because our approach does not estimate the quantities P_4 and N_4. To obtain $\widehat{N}_4$ we take the approach of setting $\widehat{N}_4$ equal to the abundance estimate that we would have obtained if no corrections a_i for availability were made. This mimics the notion that N_4 is an abundance index that does not correct for P_4.

In this case, the results of Steinhorst and Samuel (1989) and Wong (1996) apply directly. If we re-define

$$\theta_i = 1/p_i$$

and interpret the remaining notation accordingly, then our estimate of N_4 is

$$\widehat{N}_4 = \frac{1}{\widehat{E}} \sum_1^g c_i \widehat{\theta}_i$$

where

$$\widehat{\theta}_i \;=\; 1 + \exp\{-\widehat{\mu}_i - \phi_i/2\}.$$

The variance and covariance estimators for $\widehat{\theta}_i$ are

$$\widehat{\text{var}}\{\widehat{\theta}_i\} \;=\; \exp\{-2\widehat{\mu}_i - 2\phi_i\}\,(\exp\{\phi_i\} - 1)$$

and

$$\widehat{\text{cov}}\{\widehat{\theta}_i, \widehat{\theta}_j\} \;=\; \exp\{-\widehat{\mu}_i - \widehat{\mu}_j - \widetilde{\phi}_{ij}\}\,(\exp\{\phi_{ij}\} - 1)$$

(Steinhorst and Samuel, 1989). The supplementary material has further details.

George et al. (2004) define P_4 to be "the proportion of the acoustic locations directly offshore from the hydrophone array that fall within 4 km offshore from the perch" (p. 761). We compute this proportion and estimate its variance using a block bootstrap, where the blocks are chosen to be the discrete acoustic sampling periods (e.g., Givens and Hoeting 2013). Using these strategies, trend estimation proceeds as described above.

4. RESULTS

4.1. Availability

[Figure 4 about here.]

Figure 4 shows the estimated availability curve, $\widehat{f}_a(t)$. The top panel of this figure displays one point for each acoustic location in the same manner as Figure 2. The solid line in the bottom panel is the fitted availability curve on the probability scale, i.e., $\exp\{\widehat{f}_a(t)\}/(1+\exp\{\widehat{f}_a(t)\})$. The dotted lines correspond to 95% pointwise confidence intervals for each time. Averaging across time, the mean availability is 0.581; averaging across vocalizations it is 0.619.

Although this fitted curve looks quite wiggly and spans a large range of probabilities, the time span covered by this graph is 50 days, so the temporal variation in availability is not as rapid as it may appear. Further, the rate of variation matches observer impressions that migratory behavior (and ice conditions) vary every few days. The very large amount of acoustic data allows us to reliably and precisely estimate $f_a(t)$ at this temporal resolution.

4.2. Detection probabilities

The detection probability estimates of Givens et al. (2014) are described in the supplementary material. Detection probabilities were found to depend on the sighting distance (m), lead condition and group size for the ith sighting. Values ranged from about 0.3 to 0.8, and the mean was 0.495. Most standard errors were less than 0.030. See Givens et al. (2014) for further results.

4.3. Whales passing at missed times

The estimation of the effort correction for whales passing at missed times is based on the individual Horvitz-Thompson contributions h_i $(i = 1, \ldots, g)$ and their block totals H_j

$(j = 1, \ldots, 101)$. Figure 3 plots the h_i against time. Recall that the value of h_i is a number of whales, and that overall whale density and passage rate are determined by *both* the density of dots and the individual magnitudes of the h_i.

Figure 5 consolidates these data as described in Section 3.4. Figure 5 plots the estimated block counts (R_j) using one circle per block. The area of a circle is proportional to T_j (which are used as weights for fitting). The heavy curve is the spline fit for the passage rate, i.e. $f_r(t)$. Also shown with thinner (red) lines are 10 random block bootstrap pseudo-fits. A histogram of bootstrap pseudo-estimates $\widehat{E}^*$ is centered approximately on the point estimate of 0.522, and very slightly skewed right. The resulting bootstrap correction factor is $1/\widehat{E} = 1.914$ with a bootstrap standard error of 0.031.

[Figure 5 about here.]

4.4. Abundance

The point estimate of $\widetilde{N}$, without correcting for whales passing at missed times, is equal to the sum of the Horvitz-Thompson contributions, i.e., the sum of the h_i values in Figure 3. This is 8,971 whales. Adjusting for qualifying effort yields the fully corrected abundance estimate $\widehat{N} = 16,820$.

Variance calculations yield $\widehat{V}_1 = 184.85^2$, $\widehat{V}_2 = 398.73^2$ and $\widehat{\mathrm{var}}\{\widetilde{N}\} = 439.50^2$. Applying equation (6) to incorporate variability due to the effort correction yields $\widehat{\mathrm{var}}\{\widehat{N}\} = 882.84^2$. Thus, the confidence interval for the estimate is (15,176, 18,643) and the CV is 5.2%.

4.5. Trend

The estimated trend of the whale population is shown in Figure 6. The fitted growth model indicates an annual rate of increase of 3.7% with a 95% confidence interval of (2.9%, 4.6%). A pointwise 95% confidence band is also shown. This was obtained from a parametric bootstrap using the joint asymptotic distribution of the fitted parameter estimates.

[Figure 6 about here.]

5. DISCUSSION

Here we address some methodological issues and choices made during the analysis. We also examine our results in a broader context.

5.1. Exclusion of perch 2 data

Our estimator ignores the 340 whales seen only at perch 2. The reason for this is that including these sightings would require a change to the definition of detection, which in turn would greatly complicate variance estimation. Our decision does not necessarily reduce or increase the abundance estimate.

If we were to include these whales, the detection probability portion of the Horvitz-Thompson correction would need to represent $P[\text{seen from at least one perch}] = 1 - (1 - p_i)^2$ when IO is operational and $P[\text{seen at perch 1}] = p_i$ when it is not (Borchers et al., 1998). This differs from our current approach that uses only the primary perch data and the corresponding probabilities p_i. The change would introduce a quadratic function of p_i into θ_i and the denominator of the abundance estimator. For variance estimation we would need to consider expectations of exponentiations of squares of normal random variables. Compensating for this is possible; however the estimators and proofs of their asymptotic properties would be more complicated. It is not clear that the approach would make a substantial difference. The relative merits of the options are discussed by Borchers et al. (1998). We defer consideration of this alternative as a topic for possible future research.

5.2. Whales migrating outside the spatio-temporal survey region

Anecdotal evidence suggests that our estimate excludes some periods when whales passed Barrow. Although the first bowhead was seen on April 9, our analyzed season does not begin until April 13. Also, some bowhead calls were found in the acoustical recordings after the visual survey ended on June 1. The supplemental material provides further consideration of this important issue, drawing on multiple sources of evidence. We conclude that the survey covered and/or adjusted for the vast majority of the population. Nevertheless, some whales inevitably passed Barrow outside the analyzed season or area, and we recognize that this introduces a small source of downward bias in the total abundance estimate.

Our analysis explicitly accounts for whales passing during times of lapsed effort during the survey season. Other model-based methods for filling time gaps in migration counts and animals passing before/after the survey include those of Buckland and Breiwick (2002) and Mateos et al. (2012).

5.3. Bias and variance

Our approach treats $\widehat{\mathbf{\Phi}}$ and $\widehat{\mathbf{\Psi}}$ as if they are the true values of the corresponding covariance matrices. For a simpler estimation problem, the adequacy of this approximation has been simulation tested over a wide range of scenarios using the predecessor to our estimator (Wong, 1996). Generally, the results showed good bias and variance performance, even with sample sizes nearly 20 times smaller than ours. We conclude that the approximation used here has little impact on the results.

An alternative approach to variance estimation could be to apply some sort of bootstrap. This would need to respect the temporal correlation in the survey data and somehow incorporate uncertainty in detection probability estimates. The weighted likelihood estimation of detection probabilities is not easily bootstrapped (nonparametrically) due to the complex network structure of the relevant data (Givens et al., 2014).

Another source of unaccounted uncertainty is the convention of treating a Conditional whale as half a whale. The survey protocol provides little basis (e.g., confidence ratings) for a quantitative model. We therefore decided to retain the convention rather than add a new arbitrary component to our analysis.

There are several potential sources of bias worth noting. First, the counts c_i include some sightings made only with binoculars. About half of the whales were initially spotted with binoculars, at which point the observers used a theodolite to record bearing and vertical angle data from which whale location could be estimated. About 10% of the time, no theodolite sighting was obtained due to the absence of the device or an operator, or the failure to find the whale with the device despite binocular detection. Unfortunately, such 'binocular-only' data do not provide sufficiently precise estimates of range for our analyses, and the detection probability p_i cannot be estimated for these sightings. Like George et al. (2004), we do not exclude these cases. When the detection probability is not available we can scale the sighting by $1/\widehat{a}_i$ while setting $\widehat{p}_i = 1$. This corrects for the proportion of whales swimming beyond visual range while making no correction for detectability. This approach is conservative because we know that for every whale, $a_i \leq 1$ and $p_i < 1$. Therefore, the partial corrections described here will scale up the sighting less than any full correction would. For this reason, the abundance estimator will be lower than if a complete correction was available.

As noted above, a few whales pass Barrow before or after the survey season. Furthermore, baleen isotope analysis indicates that a few whales don't make the migration at all, while a few others may migrate only to Russian waters around Chukotka. As noted above, it is theoretically possible for whales to swim through the survey region entirely underwater. Although the likelihood of this is small, we do know that whales react to hunting, which is conducted sporadically some kilometers south of the perch. Also, whales may go silent or move offshore in response to noise from snow machines and planes landing in Barrow. These are all potential sources of downward bias in the abundance estimate.

The detection probability analysis is also potentially subject to sources bias. Specifically, there is likely heterogeneity in observer effects. Such unmodeled extra heterogeneity will tend to cause a downward bias in abundance estimates using standard capture-recapture abundance models (Carothers, 1973, 1979; Otis et al., 1978; Seber, 1982; Pollock et al., 1990; Hwang and Chao, 1995; Pledger and Efford, 1998; Pledger and Phillpot, 2008). Also, observers may tend to link sightings to previous sightings too often, rather than declaring the subsequent sighting to be a new whale. This would be a source of upward bias in detection probability estimates and downward bias in abundance.

There are a few sources of potential positive bias in the abundance estimate. Some apparent sightings may be something else, e.g., birds, ice, beluga or gray whales. Some whales linger in the survey area, potentially being counted twice. During periods of heavy ice whales may swim slower, again being more available for double counting. However, in such conditions they are harder to detect.

Although there are many sources of potential bias, we believe all to be relatively small. Weighing the plausibility and magnitude of these, we believe that if there is any net bias in the abundance estimate, it is downward.

5.4. Methodological considerations

Although few abundance estimation surveys would be likely to exactly mimic the bowhead case, it is clear that its individual components may be potentially useful in other surveys. A broader contribution of our work relates to the incorporation of random model-based estimated correction factors in the Horvitz-Thompson estimator and the corresponding variance. Abundance estimates that treat estimated corrections for availability and/or detection probability as fixed factors remain surprisingly common in applied statistical ecology. Our new estimator overcomes that problem. Indeed, we separately estimate those corrections from independent datasets and propagate uncertainty through to the final abundance estimate. Thus, the sampling probabilities we incorporate in the Horvitz-Thompson estimator are derived from model estimates rather than being determined by a pre-established sampling design. This general strategy is applicable to any situation where data on availability and detectability can be collected, and the derivation of the uncertainty estimate for $\widehat{N}$ in this situation is a methodological contribution of this paper.

A reviewer notes that $\int_S \widehat{f}_r(t)dt$ is an alternative abundance estimator. Although we do not pursue that idea here due to the complexities of variance estimation, we note that the corresponding point estimate would be 17,724 compared to 16,820 from our approach.

Our work has potential applications to line transect surveys as well. In our case, whales migrate past fixed perches in a mostly linear path. By changing our spatial reference, we might view the survey process as being two moving perches that linearly pass a stationary field of whales, much like a double-observer ship or airplane survey. Since the bowhead analysis is limited to 20 km off the ice edge, such a hypothetical survey would correspond to a single transect strip covering the entire population region, with model-based sampling probabilities, and there is no variance component attributable to random transect placement.

Also important is our modeling of availability and effort (via passage rate) as smooth functions to provide time-changing correction factors with appropriate uncertainty. Apart from their use in abundance estimation, these results are scientifically interesting by themselves since they describe features of bowhead migratory behavior including temporal pulses (Figure 5) and cycles of onshore/offshore passage (Figure 4).

5.5. Management implications

Indigenous hunting quotas for this population are recommended using the Bowhead Strike Limit Algorithm (SLA)–an algorithm adopted by the International Whaling Commission (IWC) after rigorous simulation testing covering a wide range of trial scenarios (International Whaling Commission, 2003). Use of this procedure would be halted if the population increase rate, both in terms of the theoretical maximum sustainable yield rate (MSYR) and/or the empirical trend estimate, is no longer believed to be in the simulation-tested range of 1% to 7% Our updated rate-of-increase estimate of 3.7% (2.9%, 4.6%) is wholly consistent with the past evidence and remains within the tested parameter space of the SLA. The most immediate management implication, therefore, is to provide continuing confidence in the SLA for setting hunting quotas.

At the time that the Bowhead SLA was adopted, the most recent abundance estimate was 10,545 in 2001 (95% CI (8,200, 13,500)). Our new estimate for 2011 is 16,820 (95% CI (15,176, 18,643)). Clearly the population size has continued to grow substantially under the levels of indigenous hunting allowed by the SLA in the last dozen years. This provides a second reason for confidence in the algorithm.

The supplementary material provides more detailed evaluation of our results in the context of other studies of these whales. The conclusion is that any biases in the 2011 survey are likely small relative to the interannual variation in abundance estimates, and the 2011 results are consistent with past findings.

Perhaps our results showing a large population abundance estimate near the naive projection, which support the status quo management approach with increasing confidence, don't seem newsworthy to a casual reader. However, aside from the statistical techniques described here, our results are actually critical for management of this population. Any whale hunting–even by indigenous communities–is extremely politically sensitive, yet such whalers have a documented subsistence and cultural need for their small hunting quota recognized by the IWC. To dampen the political firestorm, it helps to provide results from this massive, multifaceted survey project and statistical analysis showing an estimate of abundance higher than levels attained in more than a century and a strong positive rate of population growth under continuing managed hunting. There is also a pragmatic need for our efforts: the Scientific Committee of the IWC has previously recommended tapering hunting quotas to zero if an abundance estimate is not produced every ten years. Our results avert this process, which would be devastating to the native people of Alaska and Chukotka who rely on this hunt.

Rapidly changing climate and ice levels in the western Arctic contribute to a great deal of uncertainty about the future of this population, and will probably render subsistence hunting more difficult and dangerous. Bowheads thrive in heavy ice, which is becoming scarcer with passing years. This may be a significant stressor for this population. Increased oil and gas development and commercial shipping in newly opened regions may be another. On the other hand, reductions in sea ice open new potential habitat for the population such as the Northwest Passage. Our abundance and trend estimates provide benchmarks by which to evaluate the impacts of climate change and other factors influencing bowhead habitat in the years ahead.

Additional information and supplementary material for this article are available online at the journal's website.

ACKNOWLEDGMENTS

This work was supported by the National Oceanic and Atmospheric Administration (via a major grant through the Alaska Eskimo Whaling Commission (AEWC)) and the North Slope Borough (Alaska) Department of Wildlife Management. We thank the many observers, matchers, and other colleagues who conducted this challenging survey with amazing tenacity and professionalism often under dangerous field conditions. We also acknowledge the cooperation of the AEWC and thank BP Alaska for providing additional funding for field logistics. We also thank the whale hunters of Barrow who supported our studies and allowed us to conduct survey operations near their camps on the sea ice. Bailey Fosdick is thanked for helpful discussions on availability estimation. Taqulik Hepa, Harry Brower, Dolores Vinas and Molly Spicer from the North Slope Borough and Johnny Aiken and Jessica Lefevre from

the AEWC are thanked for their organizational and administrative support. Finally, Judith Zeh is thanked for her advice and wisdom about virtually every aspect of the bowhead surveys over more than 30 years.

REFERENCES

Borchers, D., Buckland, S., Goedhart, P., Clarke, E., and Hedley, S. (1998). Horvitz-Thompson estimators for double-platform line transect surveys. *Biometrics*, 54:1221–1237.

Borchers, D., Buckland, S., and Zucchini, W. (2002). *Estimating Animal Abundance: Closed Populations*. Springer Verlag.

Buckland, S. T. and Breiwick, J. M. (2002). Estimated trends in abundance of eastern Pacific gray whales from shore counts (1967/68 to 1995/96). *Journal of Cetacean Research and Management*, 4:41–48.

Carothers, A. (1973). The effects of unequal catchability on Jolly-Seber estimates. *Biometrics*, 29:79–100.

Carothers, A. (1979). Quantifying unequal catchability and its effect on survival estimates in an actual population. *J. Animal Ecology*, 48:863–869.

Clark, C., Charif, R., Hawthorne, D., Rahaman, A., Givens, G., George, J., and Muirhead, C. (2013). Analysis of acoustic data from the spring 2011 bowhead whale census at Point Barrow, Alaska. Paper SC/65a/BRG09 presented to the Scientific Committee of the International Whaling Commission, June, 2013.

Cooke, J. G. (1996). Preliminary investigation of an RMP-based approach to the management of aboriginal subsistence whaling. Paper SC/48/AS5 presented to the IWC Scientific Committee, June, 1996.

Fieberg, J. R. (2012). Estimating population abundance using sightability models: R SightabilityModel package. *Journal of Statistical Software*, 51.

George, J., Givens, G., Suydam, R., Herreman, J., Tudor, B., DeLong, R., Mocklin, J., and Clark, C. (2013). Summary of the spring 2011 ice-based visual, aerial photo-ID, and acoustic survey of bowhead whales near Point Barrow, Alaska. Paper SC/65a/BRG11 presented to the Scientific Committee of the IWC, June 2013.

George, J., Herreman, J., Givens, G., Suydam, R., Mocklin, J., Clark, C., Tudor, B., and DeLong, R. (2012). Brief overview of the 2010 and 2011 bowhead whale abundance surveys near Point Barrow, Alaska. Paper SC/64/AWMP7 presented to the IWC Scientific Committee, June 2012.

George, J. C., Zeh, J., Suydam, R., and Clark, C. (2004). 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., Edmondson, S. L., George, J. C., Tudor, B., DeLong, R. A., and Suydam, R. (2014). Weighted likelihood recapture estimation of detection probabilities from an ice-based survey of bowhead whales. *Environmetrics*, 26:1–16.

Givens, G. H. and Hoeting, J. A. (2013). *Computational Statistics, Second Edition*. John Wiley and Sons, Inc., Hoboken, NJ. 469pp.

Horvitz, D. and Thompson, D. (1952). A generalization of sampling without replacement from a finite universe. *Journal of the American Statistical Association*, 47:663–685.

Huggins, R. (1989). On the statistical analysis of capture experiments. *Biometrika*, 76:133–140.

Hwang, W.-D. and Chao, A. (1995). Quantifying the effects of unequal catchabilities on Jolly-Seber estimates via sample coverage. *Biometrics*, 51:128–141.

International Whaling Commission (2003). Chair's report of the fifty-fourth annual meeting. Annex C. Report of the aboriginal subsistence whaling sub-committee. *Ann. Rep. Int. Whaling Comm.*, 2002:62–75.

Mateos, M., Arroyo, G. M., and Thomas, L. (2012). The development and use of a method to fill time gaps in migration counts: seabird conservation applications. *The Condor*, 114:513–522.

Otis, D. L., Burnham, K. P., White, G. C., and Anderson, D. R. (1978). Statistical inference from capture data on closed animal populations. *Wildlife Monographs*. No. 62.

Pledger, S. and Efford, M. (1998). Correction of bias due to heterogeneous capture probability in capture-recapture studies of open populations. *Biometrics*, 54:888–898.

Pledger, S. and Phillpot, P. (2008). Using mixtures to model heterogeneity in ecological capture-recapture studies. *Biometrical Journal*, 50:1022–1034.

Pollock, K., Nichols, J., Brownie, C., and Hines, J. (1990). Statistical inference for capture-recapture experiments. *Wildlife Monographs*. No. 107.

Punt, A. E. and Butterworth, D. S. (1999). On assessments of the Bering-Chukchi-Beaufort Seas stock of bowhead whales (*Balaena mysticetus*) using a Bayesian approach. *Journal of Cetacean Research and Management*, 1:53–71.

R Core Team (2015). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Seber, G. (1982). *The Estimation of Animal Abundance and Related Parameters, 2nd Edition*. Griffin, London, UK.

Steinhorst, K. R. and Samuel, M. D. (1989). Sightability adjustment methods for aerial surveys of wildlife populations. *Biometrics*, 45:415–425.

Wong, C.-N. (1996). *Population size estimation using the modified Horvitz-Thompson estimator with estimated sighting probability*. PhD thesis, Colorado State University, Department of Statistics.

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. (2006). *Generalized Additive Models: An Introduction with R*. Chapman & Hall/CRC, Boca Raton, FL.

Wood, S. N. (2011). Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *Journal of the Royal Statistical Society, Series B*, 73:3–36.

Zeh, J. and Punt, A. (2005). Updated 1978-2001 abundance estimates and their correlations for the Bering-Chukchi-Beaufort Seas stock of bowhead whales. *Journal of Cetacean Research and Management*, 7:169–175.

FIGURES

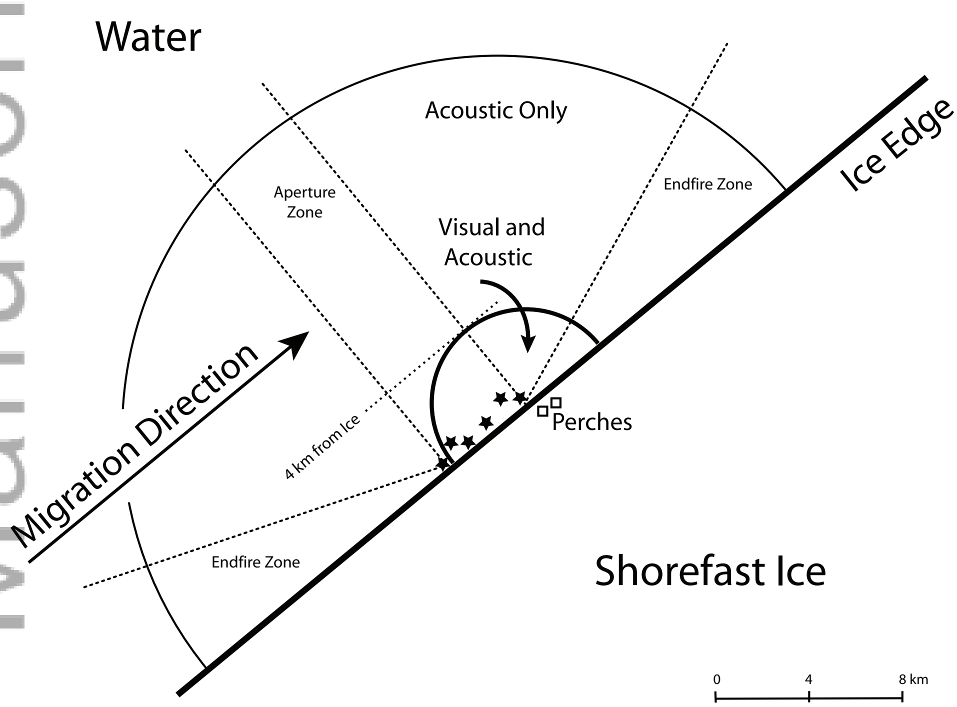

Figure 1. Layout of the 2011 visual and acoustic survey. The six acoustic recorders are stars and the two visual perches are squares. See the text for a full description. This diagram is only a sketch: for precise scale and orientation information see Clark et al. (2013).

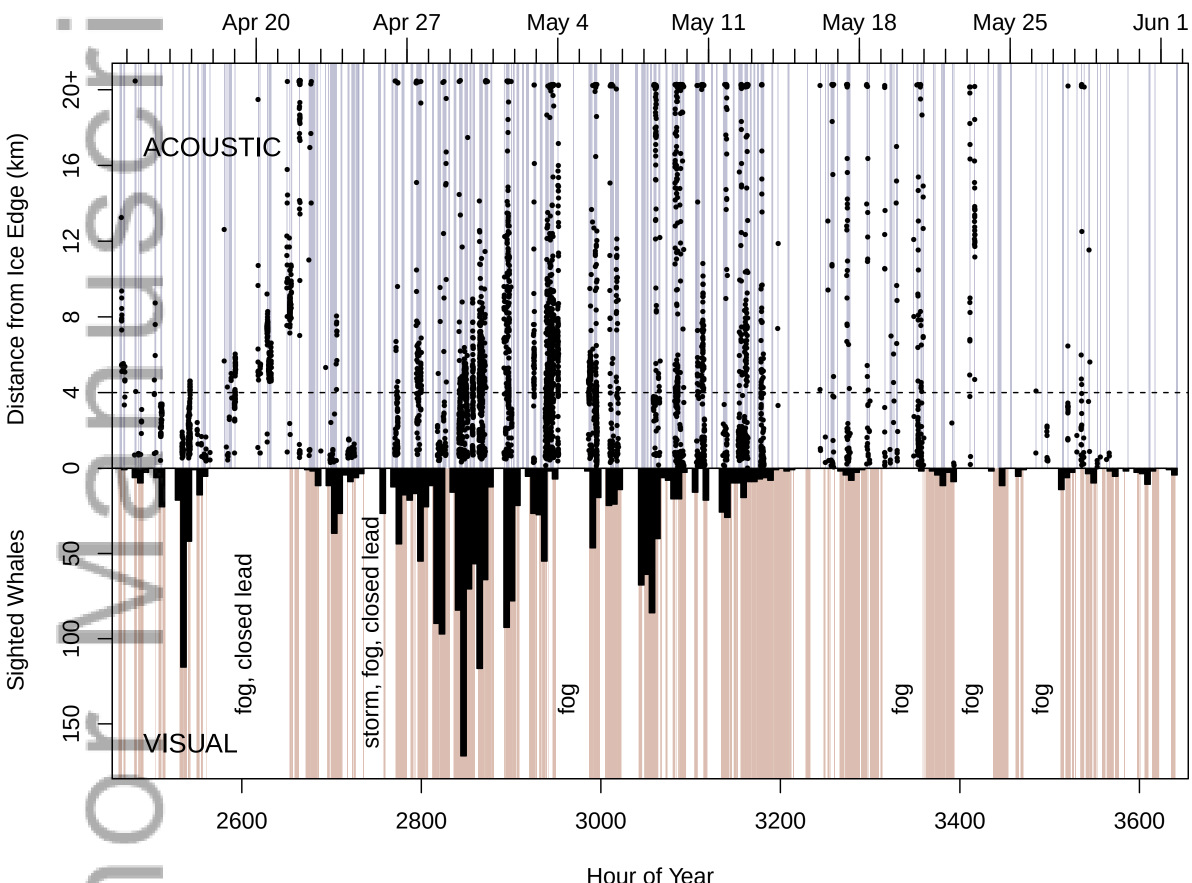

Figure 2. Summary of visual and acoustic data used in our analyses. The top portion plots individual acoustic locations by time and distance from ice edge. The bottom portion shows a (upside-down) histogram of sightings. The shaded vertical stripes correspond to time periods where data are available and white regions correspond to periods without data. See Section 2 for more details.

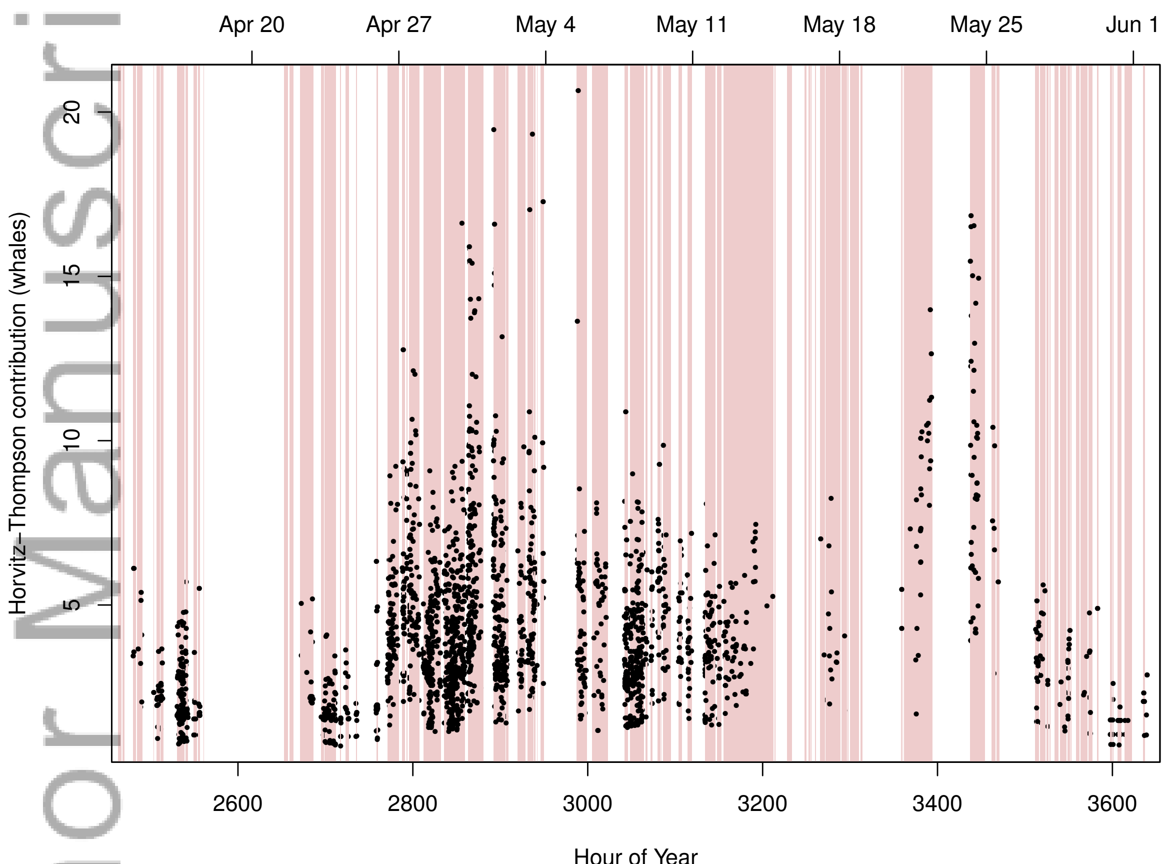

Figure 3. Horvitz-Thompson contribution, $\widehat{h}_i$, of each sighting (units are whales). The shaded bars correspond to periods of qualifying visual effort.

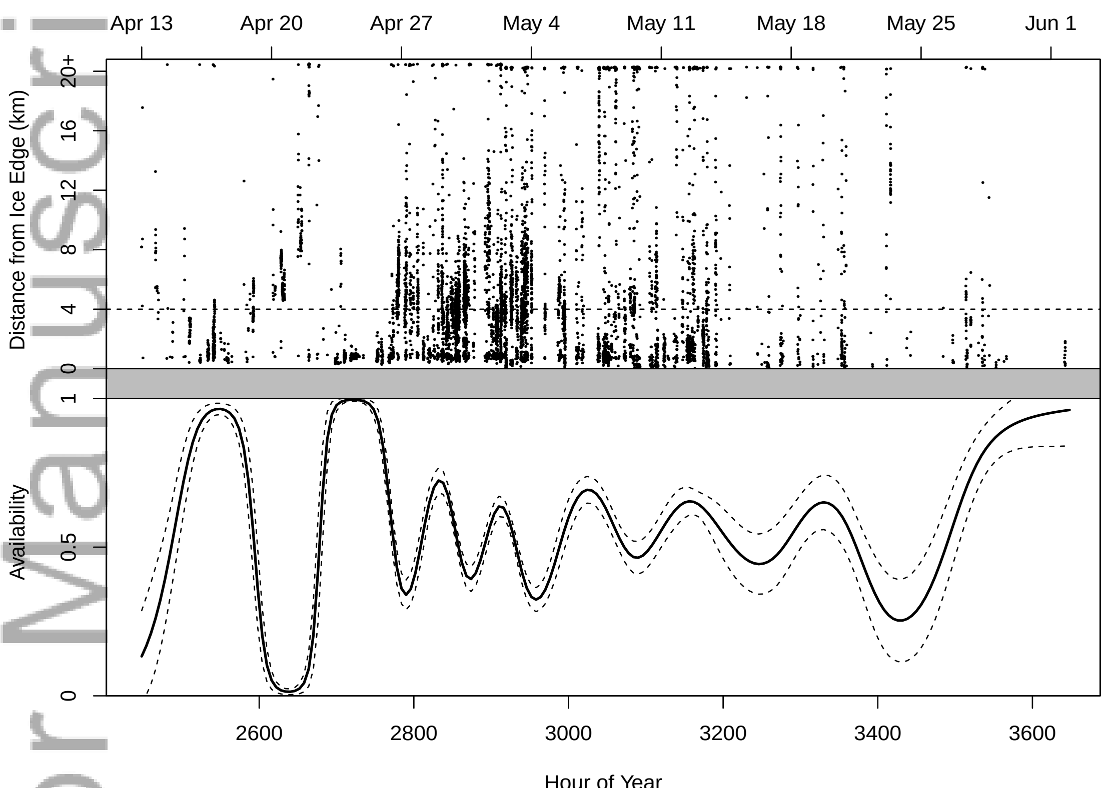

Figure 4. The top panel shows the raw acoustics data: each point represents one acoustic location at a specific time and distance from the ice edge. The bottom panel shows the estimate and pointwise 95% confidence bounds for the availability $\text{logit}^{-1}\hat{f}_a(t)$ over the course of the season. Recall that availability is defined to be the probability that a whale swims within 4 km of the ice edge and is estimated from only the acoustic data.

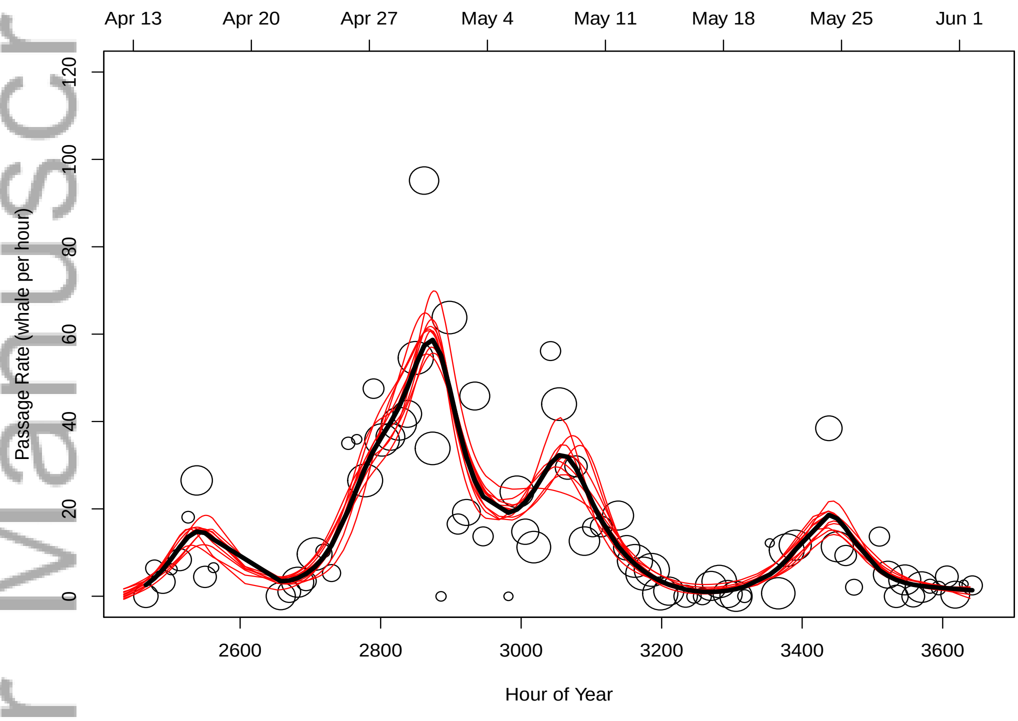

Figure 5. Estimated passage rate, $\widehat{f}_r(t)$. Block passage rates R_j (whales/hour) are shown by circles with areas proportional to T_j. The fit to these points using the gamma GAM spline is shown with the heavy line. Ten random bootstrap replicates are also shown.

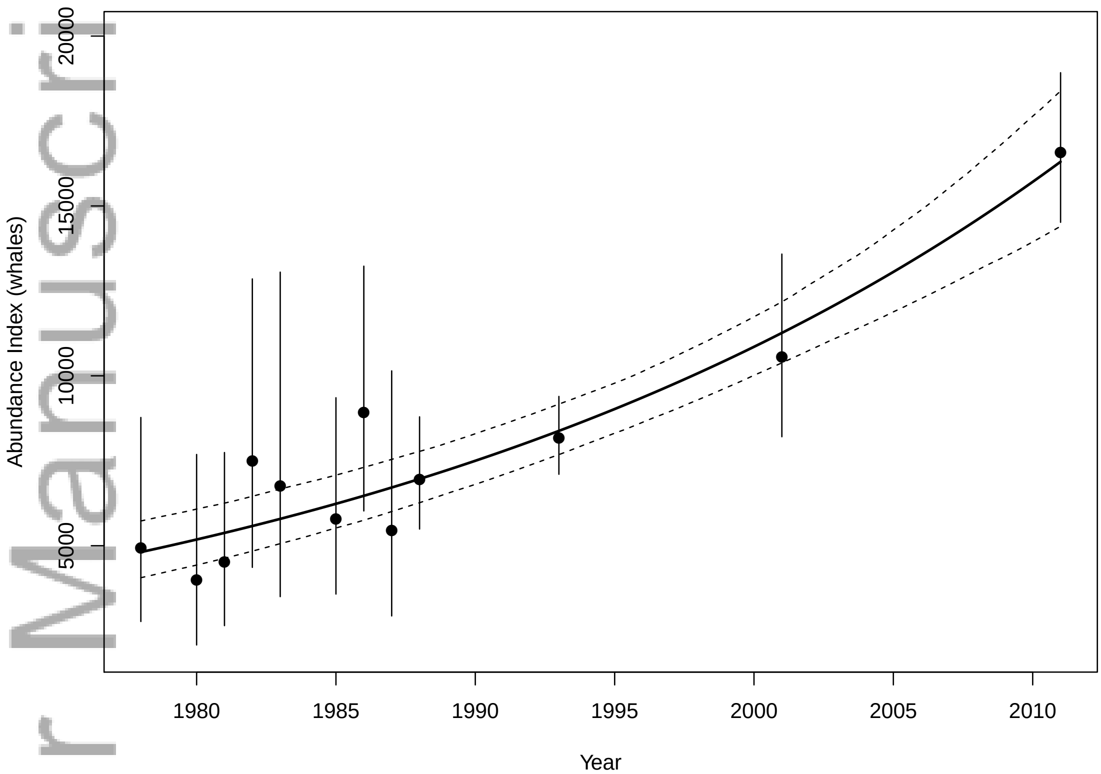

Figure 6. Estimated abundance indices, fitted curve, and pointwise 95% confidence band for the trend estimate using the time series from 1978–2011.