

## Hazard Models for Line Transect Surveys with Independent Observers

Hans J. Skaug

Institute of Marine Research, Box 1870 Nordnes, 5024 Bergen, Norway

*email:* skaug@imr.no

and

Tore Schweder

Department of Economics, Box 1095 Blindern, 0317 Oslo, Norway

**SUMMARY.** The likelihood function for data from independent observer line transect surveys is derived, and a hazard model is proposed for the situation where animals are available for detection only at discrete time points. Under the assumption that the time points of availability follow a Poisson point process, we obtain an analytical expression for the detection function. We discuss different criteria for choosing the hazard function and consider in particular two different parametric families of hazard functions. Discrete and continuous hazard models are compared and the robustness of the discrete model is investigated. Finally, the methodology is applied to data from a survey for minke whales in the northeastern Atlantic.

**KEY WORDS:** Hazard models; Independent observers; Line transect surveys.

### 1. Introduction

Line transect survey methodology is commonly used to estimate abundance of animal populations. Buckland et al. (1993) provide a comprehensive review of conventional line transect theory and practice. In conventional line transect surveys, there is a single observer (or a single team) searching for animals, and it is assumed that all animals located on the transect line are detected by the observer. This assumption is likely to be violated in several types of line transect surveys, e.g., in whale surveys. Let  $g(x)$  be the probability that an animal located at perpendicular distance  $x$  from the transect line is detected. We refer to  $g$  as the detection function. A commonly used technique for estimating  $g$  when animals on the transect line might be missed, i.e.,  $g(0) < 1$ , involves two observers surveying simultaneously the same area independently of each other. Such surveys may be viewed as two-sample mark-recapture experiments, where the animals detected by each of the two observers constitute the two different samples.

Buckland et al. (1993, p. 210) discuss a Petersen-type mark-recapture estimator of  $g(0)$ . This estimator was first used by Butterworth, Best, and Basson (1982). However, pure mark-recapture estimators of  $g(0)$  do not fully utilize the information available from independent observer surveys. Recently, Alpizar-Jara and Pollock (1996); Borchers, Zucchini, and Fewster (1998); Quang and Becker (1997); Schweder (1990); and Skaug (1997) have proposed likelihood-based methods that combine line transect techniques and mark-recapture techniques. The first three of these approaches all assume that, conditionally on the value of  $x$ , the two observers detect the

animal independently of each other. This holds for homogeneous populations, where the detection probability does not depend on factors other than  $x$ , but it does not hold in the presence of unmodelled heterogeneity in detectability, even though the observers do not communicate with each other. We shall refer to this phenomenon as dependent detection of animals. The present paper focuses on the heterogeneity resulting if animals are available for detection only in short, distinct time periods. This behavior is called discrete availability. Examples of animals with discrete availability are solitary whales being visible when they surface and birds being audible when they sing. Individuals with a high availability rate are more likely to be detected than those with a low rate. For simplicity, we refer to the cumbersome term 'event of availability' as a 'signal' given by the animal.

The concept of a hazard probability (Schweder, 1974) provides a natural framework for modelling heterogeneity caused by discrete availability. The two main components of hazard probability models are (i) the hazard probability function, which governs the observer's ability to detect animals and (ii) the signal process, which governs the availability of the animals. Hazard probability methods were first applied in animal abundance estimation by Schweder and Høst (1992). The approach was further developed by Schweder et al. (1996, 1997, 1998). The motivation for the development of the method was a need for establishing an estimate of the abundance of minke whales in the northeastern Atlantic.

In discrete availability surveys, it is natural to record the (relative) position of each detected signal, while in ordinary

line transect surveys, usually only the perpendicular distance to the animal is recorded. Of course, the perpendicular distance can be calculated using the positions of one or more of the detected signals. Thus, discrete availability data can be analyzed by taking either signal or animal as the sampling unit. Even though the analysis is performed with animal as the sampling unit, there are reasons for collecting data as individual signals. For instance, information about individual signals facilitates duplicate identification (the process of determining which animals are seen by both observers). In discarding information about individual signals, there is clearly a loss of information, so animal-based models can be expected to yield less precise estimates of the detection function than signal-based methods. On the other hand, animal-based methods are conceptually and computationally simpler. The hazard probability model introduced by Schweder et al (1991) is signal based. The approach, as further developed in Schweder et al. (1997), is rather complex since it simultaneously addresses difficulties like errors in distance measurements, errors in duplicate identification, covariate heterogeneity, and clustering in the signal process. The need for a less computationally demanding method that still does not assume independent detection of animals led Skaug (1997) to develop an animal-based hazard probability model.

Hazard models have also been developed for animals that are continuously available for detection (Buckland, 1985). Then the animals are assumed to be detected according to a hazard rate. We shall refer to such models as continuous hazard models, while the model of Skaug (1997) is referred to as a discrete hazard model. The present paper extends the model of Skaug (1997) and gives a unified treatment of discrete and continuous animal-based hazard methods for independent observer data.

The rest of the paper is organized as follows. First, an example of an independent observer survey is presented to illustrate and motivate the methodology we propose. In Section 2, the likelihood function for independent observer data is derived. In Section 3, the theory of hazard models is established. In Section 4, the robustness of the discrete hazard model is studied. In Section 5, the methodology is applied to the data described in the example below, and in Section 6, some concluding remarks are given.

#### *Northeastern Atlantic Minke Whales*

In 1995, the Norwegian government, in cooperation with the International Whaling Commission, undertook a large-scale line transect survey for minke whales in the northeastern Atlantic (see Schweder et al., 1997). Each of the 11 vessels that participated had two independent observer platforms, denoted respectively by A and B, each containing two individual observers. During the approximately 1400 hours of effective survey time, a total of 1158 minke whale sightings were reported from platforms A and B, but a considerable number of these were duplicate sightings in the sense that the same animal was reported from both A and B. We shall base the likelihood function on the set of unique animals. For several reasons, it was decided in Schweder et al. (1997) that only a subset of the 1158 sighted minke whales could be used

to estimate the detection function. Figure 1 shows a histogram of the perpendicular distances of the animals in this subset.

In the context of minke whales, a signal corresponds to a surfacing made by the whale. At each surfacing, the animal is visible only for a few seconds, and usually only the back of the animal can be seen. This makes North Atlantic minke whales difficult to detect, and it is reasonable to expect that  $g(0) < 1$ , even when weather conditions are good. We make a rough preliminary analysis to investigate this. The intuitive way of estimating  $g(0)$  is to take the animals detected by A that are close to the transect line and use the proportion of these also detected by B as an estimate of  $g(0)$ . This can be done using local linear regression with a binomial response (detected by B or not) and with  $x$  as the dependent variable. Using the S-Plus function LOESS, we find the estimate 0.44 of  $g(0)$ , with standard deviation 0.07. By changing the role of A and B, we obtain another (but not independent) estimate, 0.54, of  $g(0)$ , with corresponding standard deviation 0.10. These preliminary results indicate that  $g(0)$  is less than one and that platform A is more efficient than platform B (at least close to the transect line).

Data giving information about the surfacing process of the minke whales are needed for fitting the discrete hazard model. Sequences of dive times obtained by radio tagging of eight minke whales have been collected. Based on these data, Schweder et al. (1996) estimated a mean surfacing rate of approximately 46 surfacings per hour. They also noted an apparent cyclical pattern in the dive times series and thereby rejected the hypothesis of a Poisson signal process.

From the data described above, Schweder et al. (1997) estimated an abundance of 112,000 minke whales in the northeastern Atlantic, with a coefficient of variation of 0.1. It is not the intention of the present paper to arrive at a new abundance estimate. The purpose of studying the above example is, less ambitiously, to illustrate various aspects of our methodology on real data.

## 2. Independent Observer Data

The basic data collected in single observer line transect surveys are the perpendicular distances  $x$  from the transect line to the detected animals. The essential parameter when estimating abundance from such data is the effective strip half-width  $w$ , defined as

$$w = \int_0^\infty g(x) dx. \quad (1)$$

When  $g(0) = 1$ , we have  $w = 1/f(0)$ , where  $f$  is the probability density of  $x$ . In general, when  $g(0) < 1$ , we have

$$w = \frac{g(0)}{f(0)}. \quad (2)$$

Since the distribution of  $x$  does not depend on  $g(0)$ , data other than the perpendicular distances are needed to estimate  $w$  when  $g(0) < 1$ . We discuss estimation of  $w$  from independent observer data.

In independent observer surveys, there are two observers (or two teams, each consisting of more than one observer), denoted by A and B, with detection functions  $g_A$  and  $g_B$ , respectively. The basic data collected are the set of  $x$ 's from observer A and the set of  $x$ 's from observer B. In addition, there must be available information about which animals seen

by A were also seen by B and vice versa. The latter information is called duplicate information. A commonly used approach for estimating  $w$  is based on (2), with  $f(0)$  estimated from perpendicular distances and  $g(0)$  estimated by the Petersen-type mark-recapture estimator discussed in Buckland et al. (1993, p. 210). However, since  $w$  is defined as the integral of  $g$ , it is natural to directly fit  $g$  to the independent observer data rather than taking the detour of estimating  $f(0)$  and  $g(0)$  separately. If  $g$  is assumed to belong to a parametric family, the parameters can be estimated by maximum likelihood.

Denote the combined observer by  $A \cup B$  and let  $g_{A \cup B}(x)$  denote the probability that A, B, or both observers detect an animal located at perpendicular distance  $x$ . Our sampling units are the animals detected by  $A \cup B$ . Each observation consists of the pair  $(x, u)$ , where  $u \in \{A, B, AB\}$  indicates which of the observers made the observation ( $u = AB$  means that the animal was detected by both A and B). The joint density of  $x$  and  $u$  is

$$f(x, u) = w_{A \cup B}^{-1} \begin{cases} g_{A \cup B}(x) - g_B(x) & \text{if } u = A, \\ g_{A \cup B}(x) - g_A(x) & \text{if } u = B, \\ g_A(x) + g_B(x) - g_{A \cup B}(x) & \text{if } u = AB, \end{cases} \quad (3)$$

where  $w_{A \cup B}$  is the effective strip half-width of  $A \cup B$ . This result is established as follows. The conditional probability that  $u = A$ , given the value of  $x$ , is  $(g_{A \cup B}(x) - g_B(x))/g_{A \cup B}(x)$ . The conditional probabilities that  $u = B$  and  $u = AB$  are found similarly. Since the marginal density of  $x$  is  $g_{A \cup B}(x)/w_{A \cup B}$ , the result (3) follows.

When detections made by A and B are independent, in the sense that

$$g_{A \cup B}(x) = g_A(x) + g_B(x) - g_A(x)g_B(x), \quad (4)$$

the density (3) reduces to the density on which the likelihoods of Alpizar-Jara and Pollock (1996), Borchers et al (1998) and Quang and Becker (1997) are based. When independence does not hold, one cannot in general hope to reveal the structure of dependency solely from independent observer data. To exemplify this, assume that  $g_A = g_B = g$  and that  $g_{A \cup B}(x) = 2g(x) - k_1 g^2(x)$  for some constant  $k_1 > 0$ . When inserted into (3), this yields

$$f(x, u) = \left\{ 2 \int_0^\infty G(x) dx - k_1 g(0) \int_0^\infty G^2(x) dx \right\}^{-1} \times \begin{cases} G(x) \{1 - k_1 g(0) G(x)\}, & u = A, B, \\ k_1 g(0) G^2(x), & u = AB, \end{cases} \quad (5)$$

where  $G(x) = g(x)/g(0)$ . It is seen that  $k_1$  and  $g(0)$  only enter the expression through their product  $k_1 g(0)$ , so, from a parameter estimation perspective,  $k_1$  and  $g(0)$  are completely confounded. On the other hand, we have  $w = g(0) \int_0^\infty G(x) dx$ , such that, for the same probability distribution of  $(x, u)$ , a wide range of  $w$  values can be obtained by varying  $g(0)$  (and  $k_1$ ). In other words, the effective strip half-width is not fully determined by the distribution of  $(x, u)$ . To estimate  $w$ , we also need data on the availability of the animals.

### Parameter Estimation

Assume that  $g_A$ ,  $g_B$ , and  $g_{A \cup B}$  depend on a vector of parameters  $\beta$ . Parameters related to the process that governs the availability of the animals are not included in  $\beta$ . The log-likelihood function for data  $\{(x_i, u_i), i = 1, \dots, n\}$  is

$$l(\beta) = \sum_{i=1}^n \log \{f(x_i, u_i; \beta)\}, \quad (6)$$

with  $f$  given by (3). For reasonable choices of  $g_A$ ,  $g_B$ , and  $g_{A \cup B}$ , this will be a regular log-likelihood function and standard likelihood inference techniques can be used. Specifically, an approximation to the covariance matrix of the maximum likelihood estimator can be obtained from the observed Fisher information (Barndorff-Nielsen and Cox, 1994, p. 87).

### 3. Hazard Models

Hazard methods involve explicit modeling of the phenomenon that the observer(s) in a line transect survey is gradually approaching the animal. The animal, which may either be continuously or discretely available for detection, typically becomes increasingly detectable as the radial distance decreases. In the case of discrete availability, denote the sequence of time points of availability by  $\{T_i, i \geq 0\}$ , where, by convention,  $T_0$  is the first time point after the animal has been passed,  $T_1$  is the time point of the last given signal before passage, and so on. This is an idealized model where the observer is approaching from an infinite distance such that there are infinitely many chances of detecting the animal.

Both in the discrete and the continuous cases, we assume that the animals stay at fixed positions and that the observer is moving along the transect line at constant speed  $v$ . Let  $x$  and  $y$  be, respectively, the perpendicular and forward distance to the animal at a given time point. Consider an animal that has yet not been detected. In the discrete case, we define  $Q(x, y)$  as the conditional probability that the animal is detected, given that there is a signal at  $(x, y)$ . In the continuous case,  $Q(x, y)$  is the hazard rate at which the animal is detected, i.e., the probability of detection in a short time period of length  $h$  is approximately  $hQ(x, y)$ . In both cases,  $Q$  is called the hazard function. The hazard function is a representation of the observer's ability to detect animals at different distances. We assume that  $Q(x, y) = 0$  for  $y < 0$ , i.e., no animals are detected after they have been passed.

The discrete hazard model has a second component, i.e., the signal process, which generates  $\{T_i\}$ . We shall assume that  $\{T_i\}$  is a Poisson point process with intensity  $\alpha$ . This means that  $\alpha^{-1}$  is the average time between two consecutive signals. The Poisson assumption makes the mathematics simpler and yields a detection function of the same form as the detection function in the continuous hazard model. We refer to the discrete hazard model with  $\{T_i\}$  being a Poisson process as the Poisson-discrete model.

#### 3.1 Single Observer

In hazard models (Poisson-discrete and continuous), the detection function is given by

$$g(x) = 1 - \exp \left\{ -\frac{\alpha}{v} \int_0^\infty Q(x, y) dy \right\}, \quad (7)$$

with  $\alpha = 1$  in the continuous case. Note that  $g(0) = 1$  if and only if  $\int_0^\infty Q(0, y) dy = \infty$ . Formula (7) is derived in Buckland et al. (1993, p. 58) for the continuous case. We show it for the Poisson-discrete case.

Let  $y_0$  (negative) be the forward distance at time  $T_0$ . The detection function is given as

$$g(x) = 1 - E \left( \prod_{i \geq 1} [1 - Q\{x, y_0 + (T_i - T_0)v\}] \right),$$

where expectation is taken with respect to  $y_0$  and  $\{T_i\}$ . From the expression for the probability generating functional of the Poisson process (Cox and Isham, 1980, p. 39), it follows that  $g(x)$  is given by (7) in the Poisson-discrete case.

By definition of the Poisson process,  $T_1 - T_0, T_2 - T_1, \dots$  are independent random variables. This assumption is incompatible with any form of cyclical patterns in the intersignal times. Also, note that  $g(x)$  given by (7) depends on  $Q$  and  $\alpha$  only through the product  $\alpha Q$ . Thus, an animal that is less visible, say with hazard probability  $k_2 Q(x, y)$ ,  $k_2 < 1$ , but which is more available for detection, say with intensity  $\alpha/k_2$ , also has a detection function given by (7).

### 3.2 Independent Observers

In the discrete case, the term independent observers is taken to mean that A and B detect each signal from an animal independently of each other (given that neither A or B are aware of the animal). Thus, the hazard probability function of  $A \cup B$  is

$$Q_{A \cup B}(x, y) = Q_A(x, y) + Q_B(x, y) - Q_A(x, y)Q_B(x, y), \quad (8)$$

where  $Q_A$  and  $Q_B$  are the hazard functions of A and B, respectively. From (7), it follows that, in the Poisson-discrete model,

$$g_{A \cup B}(x) = 1 - \exp \left[ -\frac{\alpha}{v} \{q_A(x) + q_B(x) - q_{AB}(x)\} \right], \quad (9)$$

where  $q_u(x) = \int_0^\infty Q_u(x, y) dy$  for  $u = A, B$  and  $q_{AB}(x) = \int_0^\infty Q_A(x, y)Q_B(x, y) dy$ . Let  $g_{AB}(x)$  denote the probability that both A and B detect an animal that is located at distance  $x$ , i.e.,  $g_{AB}(x) = g_A(x)g_B(x) - g_{A \cup B}(x)$ . When  $g_{A \cup B}(x)$  is given by (9), we have

$$g_{AB}(x) - g_A(x)g_B(x) = \exp \left[ -\frac{\alpha}{v} \{q_A(x) + q_B(x)\} \right] \times \left[ \exp \left\{ \frac{\alpha}{v} q_{AB}(x) \right\} - 1 \right],$$

showing that A and B do not detect animals (at given distances  $x$ ) independently of each other in the discrete case and that detections are positively correlated between the two.

In the continuous case, the hazard rate function of  $A \cup B$  is  $Q_{A \cup B}(x, y) = Q_A(x, y) + Q_B(x, y)$ , and it follows that

$$g_{A \cup B}(x) = 1 - \exp \left[ -v^{-1} \{q_A(x) + q_B(x)\} \right].$$

Thus,  $g_{AB}(x) = g_A(x)g_B(x)$ , showing that A and B detect animals independently of each other in the continuous case.

### 3.3 The Hazard Function

An observer in a line transect survey searches for animals in a two-dimensional space. The hazard function  $Q$  is a quantitative representation of the observer's ability to detect an animal, viewed as a function of the location of the animal. The properties of the observer, such as skills and instructions,

can enter the model only through  $Q$ . The detection function  $g$  may be viewed as an integral projection of  $Q$  onto the  $x$ -axis through formula (7). It is not easy to visualize this projection, especially in the discrete case, since it then also involves the signal process. The hazard function, however, has a clear interpretation in terms of the actual process of approaching and detecting animals and is a natural modelling element. The difficulty in modelling the detection function is, in fact, reflected in the general absence of physical considerations surrounding the choice of parametric forms of  $g$  in the line transect literature.

The choice of model for  $Q$  should reflect physical realities as far as possible, but  $Q$  should also be chosen such that mathematical simplicity is obtained. These two concerns will often be contradictory. In the continuous case, Buckland et al. (1993, p. 59) discuss restrictions on  $Q$ . These restrictions also apply to the discrete case, but then we must have in addition that  $Q(x, y) \leq 1$ . In the following, we further discuss criteria for choosing  $Q$ .

To facilitate physical interpretation of  $Q$ , it is useful to introduce polar coordinates. Define the radial distance as  $r = (x^2 + y^2)^{1/2}$  and the sighting angle as  $\theta = \tan^{-1}(x/y)$ . The hazard function in polar coordinates is then  $Q^{(p)}(r, \theta) = Q(r \sin \theta, r \cos \theta)$ . A hazard function on the product form,  $Q^{(p)}(r, \theta) = Q^{(1)}(r)Q^{(2)}(\theta)$ , is particularly easy to interpret. If the observer spends most of his effort in the forward direction, we may choose  $Q^{(2)}$  as a decreasing function of  $\theta$  and, if the observer spends equal amounts of effort in all subsectors, we may take  $Q^{(2)}(\theta) \equiv 1$ . In Cartesian coordinates, the product form  $Q(x, y) = Q_1(x)Q_2(y)$  simplifies the mathematics. Then the detection function (7) becomes

$$g(x) = 1 - \exp \left\{ -\frac{\alpha}{v} Q_1(x)c(Q_2) \right\},$$

where  $c(Q_2) = \int_0^\infty Q_2(y) dy$ .

### 3.4 Parametric Families of Hazard Functions

This section introduces two parametric classes of hazard functions leading to expressions for  $g(x)$  and  $g_{A \cup B}(x)$  that are useful in practice. In formulas (11) and (13), the continuous case is obtained by setting  $\alpha = 1$ .

**Exponential power hazard model.** The hazard function is given by

$$Q(x, y) = \mu \exp \left( -\frac{x^\gamma + y^\gamma}{\sigma^\gamma} \right), \quad \mu, \sigma, \gamma > 0. \quad (10)$$

This class is of the product form in Cartesian coordinates. In the discrete case, we must have  $\mu \leq 1$ . Skaug (1997) used this hazard function in the special case that  $\gamma = 2$ . The interpretation of the parameters is as follows:  $\mu$  determines the level of the hazard surface,  $\sigma$  is a scale parameter, and  $\gamma$  is a shape parameter. The corresponding detection function is

$$g(x) = 1 - \exp \left[ -\alpha v^{-1} c_1 \exp \left\{ -(x/\sigma)^\gamma \right\} \right], \quad (11)$$

where  $c_1 = \gamma^{-1} \sigma \mu \Gamma(\gamma^{-1})$ , with  $\Gamma$  being the gamma function. For  $\gamma = 1$ , this detection function is a special case of (6.1) in Buckland et al. (1993, p. 184).

*Inverse power hazard model.* The hazard function is given by

$$Q(x, y) = \frac{\mu\sigma^\gamma}{(\sigma^2 + x^2 + y^2)^{\gamma/2}}, \quad \mu, \sigma > 0; \gamma > 1. \quad (12)$$

Again, we must have  $\mu \leq 1$  in the discrete case. In polar coordinates, this hazard function becomes  $Q^{(p)}(r, \theta) = \mu\{1 + (r^2/\sigma^2)\}^{-\gamma/2}$ , which does not depend on the sighting angle  $\theta$ . The interpretation of the parameters  $\mu$ ,  $\sigma$ , and  $\gamma$  is the same as for the hazard function (10). The corresponding detection function is

$$g(x) = 1 - \exp\left[-\alpha v^{-1} c_2 (x^2 + \sigma^2)^{-(\gamma-1)/2}\right], \quad (13)$$

where  $c_2 = \pi^{1/2} \mu \sigma^\gamma \Gamma\{(\gamma-1)/2\} / \{2\Gamma(\gamma/2)\}$ . This formula is obtained using (3.251.11) in Gradshteyn and Ryzhik (1994).

The detection function (13) is closely related to the Hayes–Buckland detection function (Hayes and Buckland, 1983)

$$g(x) = 1 - \exp\left\{-(x/\sigma)^{-(\gamma-1)}\right\}, \quad (14)$$

which is often used in analysis of line transect data.

The main difference between hazard functions (10) and (12) is that the latter has a heavier tail than the former. While (10) goes to zero at an exponential rate as  $r \rightarrow \infty$ , (12) goes to zero at a polynomial rate. The same is the case for the corresponding detection functions (11) and (13) as  $x \rightarrow \infty$ . The two families of models can be compared in terms of the single-observer detection functions they induce. Table 1 summarizes some key properties of the detection functions.

We end this section by asking what the advantages are of adopting a hazard model approach. In the single observer case, it lets us easily build restrictions and physical considerations into the model for the detection function  $g$ . However, in the discrete case, it is more important that the hazard approach provides the expression (9) for  $g_{A \cup B}$ , which is not based on the assumption of independent detection of animals.

#### 4. Modelling Discrete Availability

An important component of the discrete hazard model is the signal process. Because most animals will be detected only a few times (otherwise, continuous availability may be assumed as an approximation), the ordinary survey data are not well suited for gaining information about the signal process. As in the example of Section 1.1, auxiliary data may be obtained

Table 1

*Properties of the three classes of detection functions discussed in Section 3.4. A detection  $g(x)$  function is falling if it is monotonically decreasing as a function of  $x$ , it is proper when its integral  $w$  is finite, and it has a shoulder when  $g'(0) = 0$  (this is the shape criterion of Buckland et al. [1993, p. 36]).*

Detection function	$g(0)$	Falling?	Proper?	Shoulder?
Exponential power (11)	$<1$	Yes	Yes	If $\gamma > 1$
Inverse power (13)	$<1$	Yes	If $\gamma > 2$	Yes
Hayes–Buckland (14)	1	Yes	If $\gamma > 2$	If $\gamma > 1$

from specially devoted experiments. Because such auxiliary information may be both difficult and expensive to obtain, data on availability will often be collected only from a few animals in the population. The question then arises whether the behavior of these animals is representative for the whole population, and it becomes of interest to investigate the robustness of the model with respect to the assumption about the mechanism governing availability. We make a distinction between the following two levels of deviation from the model assumptions:

- (i) The average signal rate is incorrectly specified.
- (ii) The average signal rate is correctly specified, but the signal process is not a Poisson process.

Skaug (1997) studied, in a simulation experiment, the robustness of the Poisson-discrete model by generating line transect data using empirical dive-time series. He found that the resulting bias in effective strip half-width was rather small and concluded that the Poisson-discrete model is relatively robust with respect to (ii). In this section, the effect of (i) is studied. Also, we extend the Poisson-discrete model to allow heterogeneity in the signal intensity. Such a generalization is useful when data on availability indicate that different animals are available at different rates.

##### 4.1 Robustness

We study the bias in the effective strip half-width  $w$  resulting when the signal intensity  $\alpha$  is incorrectly specified. Consider the Poisson-discrete model with hazard function (10) and let  $\beta = (\mu, \sigma, \gamma)$  denote the parameter vector. The analysis is performed under the assumption that the signal intensity is  $\alpha_0$  while the true intensity is  $\alpha$  ( $\alpha \neq \alpha_0$ ). In such a setting, estimators of  $\beta$  and  $w$  will be inconsistent for the true parameter values as the sample size increases.

In general,  $w$  is a function of  $\alpha$  and  $\beta$ , so we write  $w(\beta, \alpha)$ . Let  $\hat{\beta}_0$  denote the maximum likelihood estimator of  $\beta$  based on the erroneous log-likelihood  $l(\beta; \alpha_0)$ , where  $l(\cdot; \alpha)$  is given by (6). When the number of observations becomes large,  $\hat{\beta}_0$  goes to  $\tilde{\beta}$ , defined as the value of  $\beta$  minimizing the Kullback–Liebler distance (Bickel and Doksum, 1976, p. 226) between the true density  $f(x, u; \beta, \alpha)$  and the assumed density  $f(x, u; \tilde{\beta}, \alpha_0)$ . The apparent effective strip half-width  $w_0 = w(\tilde{\beta}, \alpha_0, \alpha_0)$  has the interpretation as the estimate of  $w$  obtained under the erroneous assumption that the signal intensity is  $\alpha_0$ .

As a numerical example, we have chosen  $\mu = 0.3$ ,  $\sigma = 700$ ,  $\gamma = 2$ , and  $\alpha_0 = 0.0129$ . These are rough estimates for North Atlantic minke whales (see Skaug, 1997). When plotting the relative error in the apparent effective strip half-width,  $r_w(\alpha) = \{w_0(\alpha) - w(\alpha)\}/w(\alpha)$ , against the relative error in the assumed signal intensity,  $r(\alpha) = (\alpha_0 - \alpha)/\alpha$ , the near linear relation  $r_w(\alpha) = 0.29r(\alpha)$  emerges when  $\alpha$  ranges between 0.006 and 0.019. There is a considerable dampening effect of the error in  $\alpha$  on the error in  $w$ , so the model seems fairly robust to failure of model assumptions of type (i).

##### 4.2 Random Heterogeneity in Availability

The signal intensity  $\alpha$  may vary between animals in the population, and it may vary over time for a given individual. For instance, feeding animals may have a different signal rate than nonfeeding animals. For a randomly sampled animal, it is natural to view  $\alpha$  as a random variable. The population

detection function, obtained by taking the expectation with respect to  $\alpha$  in (7), is

$$g(x) = 1 - L \left\{ -v^{-1} \int_0^\infty Q(x, y) dy \right\},$$

where  $L$  is the Laplace transform of the distribution of  $\alpha$ . In the special case that  $\alpha$  has a gamma distribution with parameters  $\tau$  and  $\kappa$ , such that  $E(\alpha) = \kappa/\tau$  and  $\text{var}(\alpha) = \kappa/\tau^2$ , we obtain

$$g(x) = 1 - \left\{ 1 + (\tau v)^{-1} \int_0^\infty Q(x, y) dy \right\}^{-\kappa}.$$

## 5. Northeastern Atlantic minke whales

The data from the survey were described in Section 1. Three models were fitted to the data. The models are

- (1) The exponential power hazard model, with  $g_{AUB}$  given by (9) and  $Q$  by (10).
- (2) The generalized Hayes–Buckland model with independent detection of animals, with  $g_{AUB}$  given by (4) and with  $g(x) = \mu g_{HB}(x)$ , where  $0 < \mu \leq 1$  and  $g_{HB}$  is given by (14).
- (3) The exponential power hazard model with independent detection of animals, with  $g$  given by (11) and  $g_{AUB}$  given by (4).

The details of how the analysis was carried out are as follows:

- The perpendicular distance of animals was calculated from the first detected surfacing.
- It was assumed that platform A and B had the same hazard function, although the preliminary analysis made in Section 1.1 indicated the contrary.
- The same set of duplicate whales that was identified by Schweder et al. (1997) was used. Possible errors in the duplicate identification were ignored.
- Measurement error in observed radial distances was ignored.

The parameter estimates for the three models, with corresponding estimates of  $g(0)$  and  $w$ , are given in Table 2. Standard deviations of the estimators  $\hat{\mu}$ ,  $\hat{\sigma}$ , and  $\hat{\gamma}$  were obtained from the observed Fisher information. The standard deviations of  $\hat{g}(0)$  and  $\hat{w}$  were found by simulation, where in each simulation loop  $(\hat{\mu}, \hat{\sigma}, \hat{\gamma})$  was drawn from its estimated sampling distribution and, for these parameter values,  $g(0)$  and  $w$  were calculated.

It is seen that there is a 30% difference in  $w$  between Model 1 and Model 2. From the discussion following equation (5), it is clear that two models that fit data equally well can yield widely different values of  $w$ . It is thus natural to ask if the observed difference in  $w$  is caused by the assumption about independent detection of animals made for Model 2 or if it is simply due to the fact that one of the models fits data better than the other. To clarify the picture, we look at the results for Model 3. It is seen that the estimate of  $w$  is almost identical to that of Model 2. This, together with the fact that Model 1 and Model 3 have an almost identical fit to data, indicate that the 30% difference in  $w$  is due to the fact that Model 2 assumes independent detection of animals.

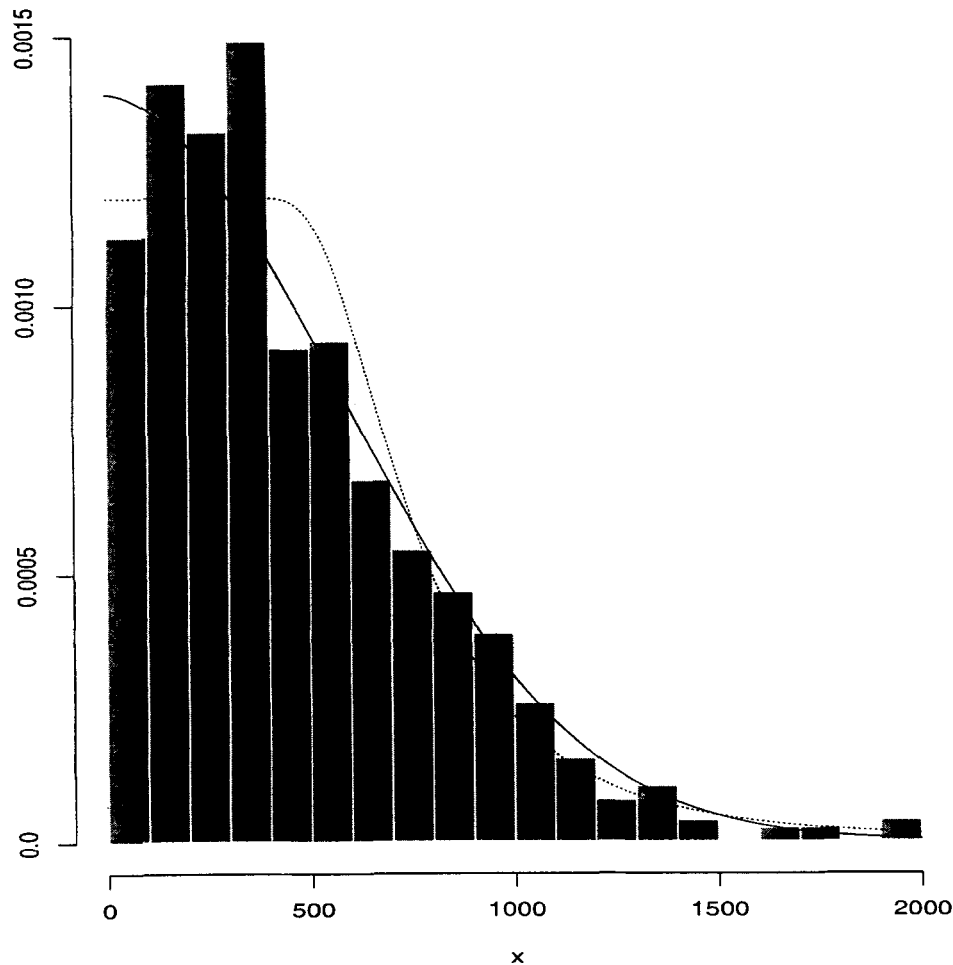
Goodness-of-fit should be assessed in at least two respects: the fit to the observed perpendicular distances and the fit to the mark–recapture part of data. Figure 1 shows a histogram of  $x_1, \dots, x_n$  together with the estimated probability density of  $x$  for Models 1 and 2. The density function of Model 3 is almost indistinguishable from the density of Model 1 and is hence not plotted. It seems from the histogram that the density of sightings is lower in the interval  $[0, 100]$  than it is in the interval  $[100, 400]$ . We have no good explanation for this phenomenon. Because both the fitted densities are monotonously decreasing as functions of  $x$ , neither of them fit data well at  $x = 0$ . A Kolmogorov–Smirnov goodness-of-fit test gave a p-value of 0.20 for Model 1, which indicates that the overall fit is acceptable. To assess the fit to the data  $u_1, \dots, u_n$ , we compare  $P(u = AB) = 2w/w_{AUB} - 1$  to the empirical frequency of  $u = AB$ . We find this empirical frequency to be 0.170, with a standard deviation of 0.01. The proportion of observations with  $u = AB$  as predicted by each of the models is 0.171 for Model 1, 0.167 for Model 2, and 0.169 for Model 3. These differences are so small that they cannot explain the 30% difference in  $w$  discussed above.

## 6. Discussion and Conclusion

The main proposal of this paper is a hazard method for analysis of data from independent observer line transect surveys where the animals are available for detection only at discrete time points. The discrete availability has two important consequences for the analysis. First, from real world considerations, we expect that  $g(0) < 1$  due to the discrete availability. For both of the hazard models (10) and (12) we have  $g(0) < 1$ , but the hazard framework also allows models with  $g(0) = 1$ , such as the Hayes–Buckland model. If one, from practical considerations, expects that  $g(0) = 1$  despite the discrete availability, conventional single observer surveys should be used instead of independent observer surveys. The second consequence of discrete availability is that the detections made by the independent observers are, somewhat paradoxically, not statistically independent. If not accounted for in the analysis, both the fact that  $g(0) < 1$  and the fact that detections are dependent will bias the estimate of the effective strip half-width.

Although it is not easy to separate these two sources of bias, we will try to assess their individual magnitudes in the case of the survey for northeastern Atlantic minke whales that was analyzed in Section 5. If one adopts the formula  $g(0)/f(0)$  for the effective strip half-width, the relative error made by assuming  $g(0) = 1$  is given as  $g(0)^{-1} - 1$ . For the three models of Table 2, we find that these numbers are 2.1, 1.9, and 1.4, respectively. On the other hand, the effect of assuming independent detection of animals was found in Section 5 to be an approximately 30% increase in  $w$ . Thus, an assumption that  $g(0) = 1$  yields a much larger error than an assumption of independent detection of animals in the case of North Atlantic minke whales.

An important goal in developing the present method has been mathematical and computational simplicity. Simplicity is ensured by the assumption that the time points of availability of the animals are governed by a Poisson point process. Although it was allowed in Section 4 that be different animals



**Figure 1.** Histogram of perpendicular distances for the minke whale data. Also plotted are the estimated perpendicular distance densities of the exponential power hazard model (solid line) and the generalized Hayes-Buckland model (dotted line).

**Table 2**

*Parameters estimated from the northeastern Atlantic minke whale data. Model 1: The exponential power hazard model. Model 2: The generalized Hayes-Buckland model with independent detection of animals. Model 3: The exponential power hazard model with independent detection of animals. Note that  $\mu$ ,  $\sigma$ , and  $\gamma$  represent different parameters in Model 2 than in the two other models. The effective strip half-width is given in meters.*

	$\mu$	$\sigma$	$\gamma$	$g(0)$	$w$
<b>Model 1</b>					
Estimate	0.250	700	1.70	0.320	210
Standard deviation	0.033	42	0.13	0.025	14
<b>Model 2</b>					
Estimate	0.350	660	5.10	0.350	280
Standard deviation	0.024	32	0.30	0.024	20
<b>Model 3</b>					
Estimate	0.360	680	1.70	0.420	280
Standard deviation	0.040	41	0.13	0.024	17

can available at different rates, the method relies heavily on the Poisson assumption. Sensitivity with respect to the Poisson assumption was discussed in Section 4. The model seems to be robust in this regard, but we have of course only investigated one parameter setting. Finally, we have not discussed estimation of covariate effects. Our aim has been to derive the basic components of the discrete hazard model. For North Atlantic minke whales, the effect of covariates probably is of at least the same magnitude as the effect of dependent detection.

## 7. Software

The first author has developed a set of S-Plus routines for fitting the Poisson-discrete hazard model. In the present version of the program, only the hazard function (10) is implemented and it is required that  $g_A = g_B$ . The S-Plus function NLMINB is used to perform the optimization of the likelihood function. The program calculates analytical derivatives of the log-likelihood function in order to speed up convergence and to increase numerical stability. Fitting the Poisson-discrete model to the data set of Section 5 took 15 seconds on a Sun Ultra Sparc. The software can be requested by sending email to HJS (skaug@imr.no).

## ACKNOWLEDGEMENTS

We are grateful to the editor, the associate editor, and three anonymous referees for comments that helped us improve the paper considerably.

## RÉSUMÉ

La fonction de vraisemblance pour des données de transects en lignes obtenues par des observateurs indépendants est construite et un modèle de risque est proposé pour le cas d'animaux détectables seulement à des dates discrètes. Sous l'hypothèse que les dates de détectabilité suivent un processus de Poisson, nous obtenons une expression analytique de la fonction de détection. Nous discutons différents critères de choix de la fonction de risque, en considérant en particulier deux familles paramétriques. Des modèles de risque discrets et continus sont comparés et la robustesse du modèle discret est étudiée. Enfin, la méthode est appliquée à des données de suivi de Petit Rorqual dans le nord-est de l'Atlantique.

## REFERENCES

- Alpizar-Jara, R. and Pollock, K. H. (1996). A combination line transect and capture recapture sampling model for multiple observers in aerial surveys. *Environmental and Ecological Statistics* **3**, 311–327.
- Barndorff-Nielsen, O. and Cox, D. R. (1994). *Inference and Asymptotics*. London: Chapman and Hall.
- Bickel, P. J. and Doksum, K. A. (1976). *Mathematical Statistics*. Oakland: Holden-Day.
- Borchers, D. L., Zucchini, W., and Fewster, R. M. (1998). Mark-recapture models for line transect surveys. *Biometrics*, in press.
- Buckland, S., Anderson, D., Burnham, K., and Laake, J. (1993). *Distance Sampling; Estimating Abundance of Biological Populations*. London: Chapman and Hall.
- Buckland, S. T. (1985). Perpendicular distance models for line transect sampling. *Biometrics* **41**, 177–195.
- Butterworth, D. S., Best, P. B., and Basson, M. (1982). Results of analysis of sighting experiments carried out during the 1980/81 southern hemisphere minke whale assessment cruise. *Report of the International Whaling Commission* **32**, 819–834.
- Cox, D. R. and Isham, V. (1980). *Point Processes*. London: Chapman and Hall.
- Gradshteyn, I. and Ryzhik, I. (1994). *Tables of Integrals, Series, and Products*, 5th edition. San Diego: Academic Press.
- Hayes, R. J. and Buckland, S. T. (1983). Radial-distance models for the line-transect method. *Biometrics* **39**, 29–42.
- Quang, P. X. and Becker, E. F. (1997). Combining line transect and double count sampling techniques for aerial surveys. *Journal of Agricultural, Biological, and Environmental Statistics* **2**, 230–242.
- Schweder, T. (1974). Transformation of point processes: Application to animal sighting and catch problems, with special emphasis on whales. Ph.D. thesis, University of California, Berkeley.
- Schweder, T. (1990). Independent observer experiments to estimate the detection function in line transect surveys of whales. *Report of the International Whaling Commission* **40**, 349–355.
- Schweder, T., Øien, N., and Høst, G. (1991). Estimates of the detection probability for shipboard surveys of northeastern Atlantic minke whales, based on a parallel ship experiment. *Report of the International Whaling Commission* **41**, 417–432.
- Schweder, T., Hagen, G., Helgeland, J., and Koppervik, I. (1996). Abundance estimation of northeastern Atlantic minke whales. *Report of the International Whaling Commission* **46**, 391–405.
- Schweder, T., Skaug, H. J., Dimakos, X., Langaas, M., and Øien, N. (1997). Abundance estimates for northeastern Atlantic minke whales. Estimates for 1989 and 1995. *Report of the International Whaling Commission* **47**, 453–484.
- Schweder, T., Skaug, H. J., Langaas, M., and Dimakos, X. (1998). Simulated likelihood analysis of line transect data. *Biometrics*, in press.
- Skaug, H. J. (1997). Perpendicular distance line transect methods based on integrated spatial hazard probability models. *Report of the International Whaling Commission* **47**, 493–498.

Received March 1997. Revised April 1998.

Accepted May 1998.