Technical Reports of the Institute of Cetacean Research (2019) pp. 14–20

Technical Report (not peer reviewed)

A note on *g*(0) estimates derived from vessel-based sighting surveys

Megumi TAKAHASHI*

Institute of Cetacean Research, 4–5 Toyomi-cho, Chuo-ku, Tokyo 104–0055, Japan

**Contact e-mail: takahashi@cetacean.jp*

ABSTRACT

This article introduces the concept of *g*(0), the probability of detection of animals at zero distance from the trackline, in the development of abundance estimates of large whales. In the conventional line transect method, it is assumed that *g*(0)=1, meaning that all animals on the trackline are detected. However, this assumption is sometimes violated, which causes underestimation of the abundance. Also this article provides overviews of the methods used to estimate *g*(0) for large baleen whales, and of the current work of the Institute of Cetacean Research to estimate *g*(0) based on dedicated sighting surveys in the Antarctic.

INTRODUCTION

Abundance estimates of animals can be based on total counts (census), sample counts, index counts, among others. The method producing the most accurate estimate is probably the total census of animals in a given area. This is possible when the area is small, the visibility is good, the target species is easily detectable, and when there is no emigration and immigration. However, these conditions are difficult to meet for most of the large whale species. In such cases, the sample counts approach is more appropriate. Under this approach, the abundance estimate is extrapolated from data obtained from a part of the total research area.

Distance Sampling (DS) methods (Buckland *et al.*, 2001) are widely used for estimating whale abundance in large areas. The line transect method is one of those methods, in which detailed surveys are conducted along a trackline that covers part of the entire survey area (Figure 1A). The detection probability of whales is not constant but decreases with the distance from the trackline (Figure 1B). Hence, in conventional line transect methods it is assumed that the detection probability on the trackline, *g*(0), is equal to 1. This means that all whales on the trackline are detected (Buckland *et al.*, 2001).

However, whale schools are not always detected even if they are on the trackline because whales dive and because observers could miss the schools on the trackline. In such cases a *g*(0) estimate is required for correction of abundance estimates.

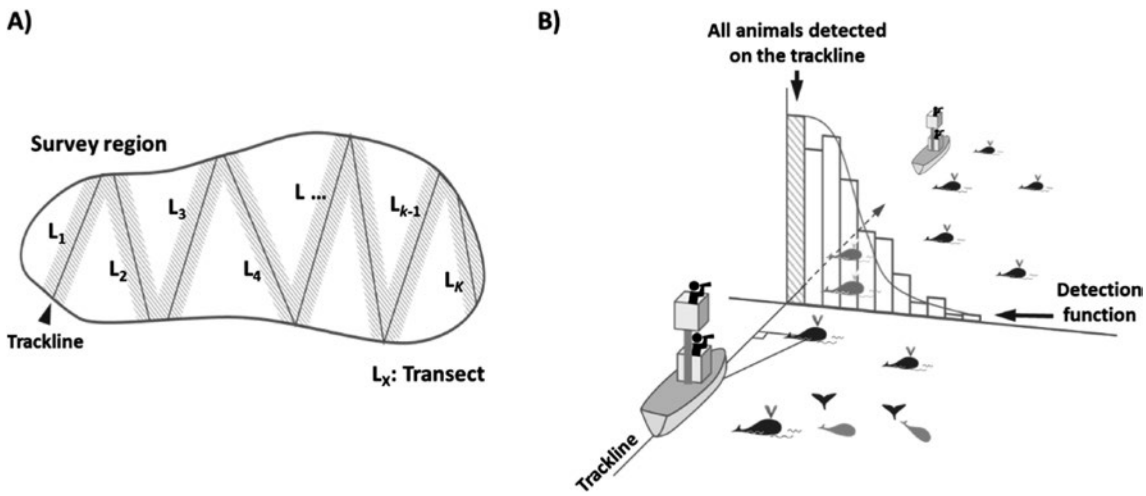

Figure 1. Diagram showing the zigzag tracklines in part of the total research area under the line transect survey (A); and the pattern of detection probability in relation to perpendicular distances (B).

The objective of this article is to introduce the concept of *g*(0) in the development of abundance estimates of large whales. Also this article provides overviews of the methods used to estimate *g*(0) for large baleen whales, and of the current work of the Institute of Cetacean Research (ICR) to estimate *g*(0) based on dedicated sighting surveys in the Antarctic.

AVAILABILITY AND PERCEPTION BIASES

Under the conventional DS methods (Buckland *et al.*, 1993; 2001), *g*(0) is assumed as equal to 1. However, when this assumption is violated, conventional DS estimators result in density and abundance being underestimated. In the case of whales, they are not always detected because they spend time underwater. This is known as 'availability bias' (Figure 2). This bias is potentially the largest problem for long-diving species that spend little time at the surface when the boat or ship surveys are conducted. For diving species, it may be necessary to model diving behavior to estimate the availability bias (e.g. Barlow, 1999).

Also, the observers could miss some whales available at the surface for a variety of reasons including visibility such as glare and mist, fatigue or moments of inattention. This is known as 'perception bias' (Figure 2). This bias is potentially largest for species that occur as single animals or in small groups and do not show much of their body when surfacing, such as Hector's dolphin and minke whales (Dawson *et al.*, 2004).

Both availability and perception biases may reduce the detection probability. In many cases, there is a combined effect of these on estimates of abundance and density for target species. The term 'visibility bias' is used generically to refer to either or both types of biases (Laake and Borchers, 2004).

If the assumption of *g*(0) = 1 is violated then the abundance estimates of the target species will be underestimated. For example, Marsh and Sinclair (1989) estimated that 83.3% of dugons (*Dugong dugon*) were beneath the water surface and unavailable for detection, while the observer team only missed 2–17% of the visible dugongs within a 200 m strip. Laake and Borchers (2004) noted that the availability bias should not be ignored if it occurs, because it can be a substantial source of bias that may be much larger than perception bias.

Therefore estimates of *g*(0) for cases of availability and perception biases are important and data required for such estimates should be collected.

DATA REQUIRED FROM THE FIELD TO ESTIMATE *g*(0) IN CASE OF LARGE WHALES

The ICR conducts sighting surveys based on the distance sampling for estimating abundance of large whales. In general, the surveys follow the Requirements and Guidelines for Conducting Surveys and Analysing Data within the Revised Management Scheme of the International Whaling Commission Scientific Committee (IWC SC) (IWC, 2012). Details of the procedures are found in Hakamada and Matsuoka (2017).

These surveys have been implemented on pre-determined zigzag tracklines. The start point is randomly selected for each survey. The design takes care not to follow physical features such as isobaths that may be correlated with whale distribution and their migration. For each whale school sighted, the species is identified and other data are recorded such as sighting position, school

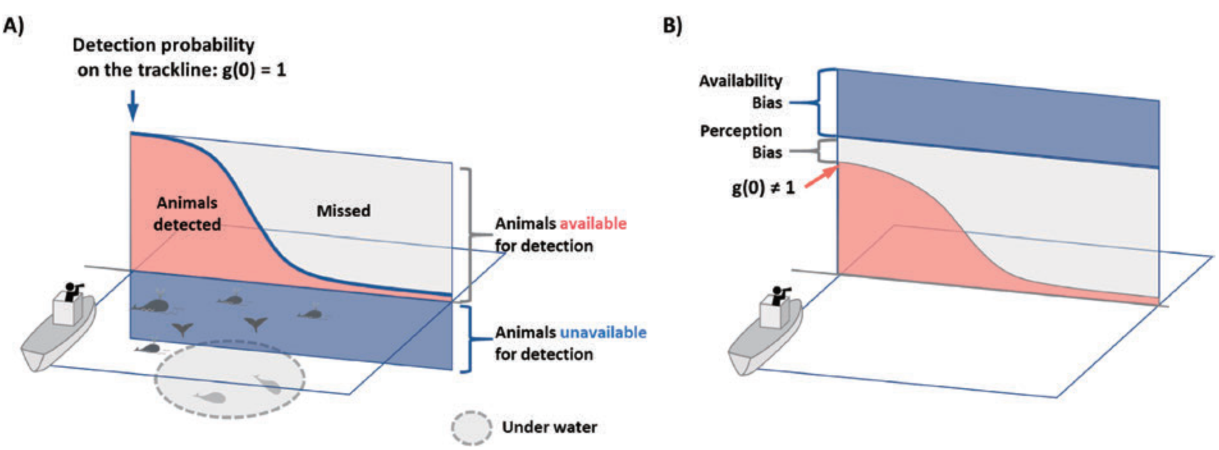

Figure 2. Diagram showing the detection probability with assumed *g*(0) = 1 (A); and the case of considering availability and perception biases when *g*(0) is different from 1 (B) (modified from https://workshops.distancesampling.org/stand-intermed-2018/slides/mrds1-g0.pdf).

size, weather information, etc.

Sighting data are collected from two platforms available in the same vessel, which allows accounting for animals missed on the trackline. Sighting surveys are conducted by using a top barrel platform (TOP) and an independent observer platform (IOP). Sighting surveys are conducted from two platforms independently, following established procedure protocols (Butterworth and Borchers, 1988; Palka, 1995; Matsuoka *et al.*, 2003; IWC, 2012).

Under the IO mode, sighting surveys are conducted from the TOP and the IOP. Personnel at the upper bridge of the vessels record the sighting information from the two platforms and determine if the same sighting is made from the two platforms or if some sightings are missed from one of the platforms.

Figure 3 shows the research vessel *Kaiyo-Maru* No.7 (KY7) equipped with TOP, IOP and upper bridge. The vessel has instruments allowing contact between TOP and upper bridge, and between IOP and upper bridge.

During the survey, one or more observers are always in the TOP, and one observer is in the IOP. Other observers are in the upper bridge getting sighting information from the two platforms. Observers in the top barrel and IOP report the sightings to the upper bridge observers, but that information is not interchanged between TOP and IOP.

Figure 4 shows a diagram of double-platform sighting survey and judgment of duplicate status. Here, a single whale appears three times, which is detected by platform A (TOP). On the other hand, platform B (IOP) detects the same whale only during the last surfacing (A3). Observers in the upper bridge receive sighting information (e.g. distance, angle, species and school size) from platform A at the first sighting (A1) and track this single whale until platform B observed it. They make a judgment concerning the duplicate status (same whale sighted by the two platforms).

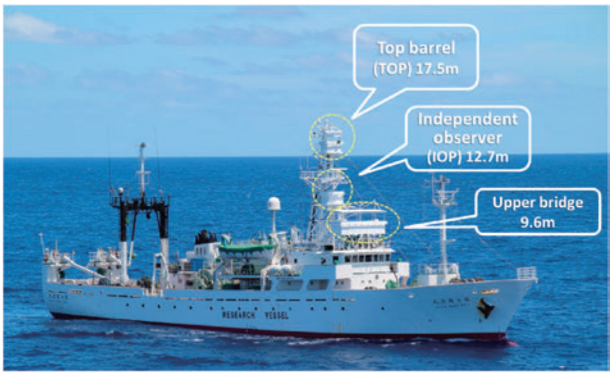

Figure 3. Survey vessel equipped with three platforms for conducting surveys required for *g*(0) estimate.

ANALYTICAL APPROACHES TO ESTIMATE *g*(0)

MRDS method

The method called 'Mark-Recapture Distance Sampling' (MRDS) is one of the methods employed to estimate *g*(0), which consists of two models: a multiple covariate Mark-Recapture (MR) model (for estimating observer detection rates) and a multiple covariate Distance Sampling (DS) model (for estimating the variation of detection probabilities with distance from the vessel) (Buckland and Turnock, 1992; Alpizar-jara and Pollock, 1996; Quang and Becker, 1997; Borchers *et al.*, 1998; Innes *et al.*, 2002; Laake and Borchers, 2004; Borchers *et al.*, 2006, Burt *et al.*, 2014). The MRDS method deals only with perception bias.

The probability that either or both observers detect a school (or animals) is given by:

$$\hat{P}.(y,z) = \hat{P}_1(y,z) + \hat{P}_2(y,z) - \hat{P}_1(y,z)\hat{P}_2(y,z),$$

where y is the perpendicular distance corrected and z denotes the covariate. Although observers 1 and 2 are considered independent from each other under the IO mode survey, the detection probability of observers can be correlated because of factors such as school size. This heterogeneity is denoted in the probabilities of detection using a logistic form for the detection function:

$$P_{l|3-l}(y,z) = \frac{\exp\left(\beta_0 + \beta_1 y + \sum_{k=1}^{K} \beta_{k+1} z_k\right)}{1 + \exp\left(\beta_0 + \beta_1 y + \sum_{k=1}^{K} \beta_{k+1} z_k\right)},$$

where, l can take the values 1 or 2 to represent the observers, β_0, $\beta_1 \cdots \beta_K$ represent the parameters to be esti-

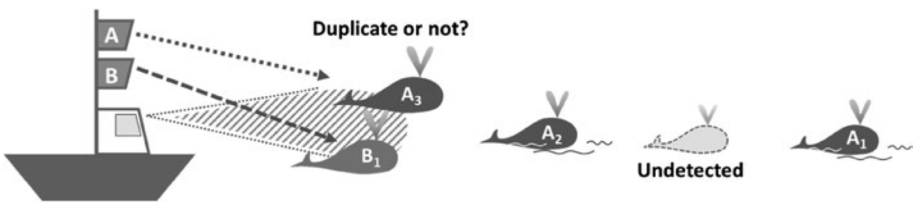

Figure 4. Diagram of the double-platform sighting survey conducted by the ICR.

mated, and K is the number of covariates. When perpendicular distance differed between duplicates, the average distance of the duplicate pair are used.

On the other hand, the detection probability away from the trackline is described and estimated by the DS model as follows:

$$\text{Half-normal: } f(y,z) = \exp\left(-\frac{y^2}{2\sigma(z)^2}\right),$$

$$\text{Hazard-rate: } f(y,z) = 1-\exp\left[\left(\frac{-y}{\sigma(z)}\right)^{-b}\right],$$

where σ and b are the parameters of each functional form.

Both models can include covariates (e.g. school size, Beaufort, etc.). In addition, the likelihoods can be treated as MR model and DS model separately or as combination of those models. The best model usually is selected using the Akaike Information Criterion (AIC; Akaike, 1973).

The MRDS method do not consider the availability bias. Hence, the availability bias should be accounted by analyses of data from additional experiments (e.g. Heide-Jørgensen *et al.*, 2010).

Hazard probability model: ʻOKʼ method

This method deals with both availability and perception biases. The method, called the ʻOKʼ method, was developed by Okamura *et al.* (2003) by extending the model of Skaug and Schweder (1999), and by combining the merits of the models of Schweder *et al.* (1997) and Cooke (2001). Okamura *et al.* (2003) expanded the concept of the hazard probability model (Skaug and Schweder, 1999) to avoid the use of external information for considering diving behavior (the conventional and simple hazard probability models require an estimate of the surfacing rate based on external data).

The hazard probability function $Q(x,y)$ for an observer u (u=A or B) is assumed as:

$$Q_u(x,y) = \mu_u \exp\left\{-\left(\frac{x}{\sigma_u}\right)^{\gamma_1} - \left(\frac{y}{\sigma_u}\right)^{\gamma_2}\right\},$$

where $0<\mu_u \leq 1$, $\sigma_u>0$, $\gamma_1, \gamma_2>0$, μ_u is the level parameter of the hazard probability function, σ_u is the scale parameter, and γ_1 and γ_2 are the shape parameters (Skaug and Schweder, 1999). The corresponding detection function from this hazard function is explicitly expressed as:

$$g_u(x) = 1-\exp\left\{-\left(\frac{\lambda}{v}\right)\int_0^\infty Q_u(x,y)\,dy\right\}$$
$$= 1-\exp\left[-\lambda v^{-1} c_2^u \exp\left\{-\left(\frac{x}{\sigma_u}\right)^{\gamma_1}\right\}\right],$$

where λ is surfacing intensity, v is constant vessel speed and $c_2^u = \sigma_u \mu_u \gamma_2^{-1} \Gamma(\gamma_2^{-1})$, with Γ being the gamma function.

The ʻOKʼ method is a more general and comprehensive method that integrates the merits of previous hazard probability models.

ONGOING STUDIES IN THE ANTARCTIC

The ʻOKʼ method explained above (Okamura *et al.*, 2003; Okamura and Kitakado, 2012) was used previously for estimating Antarctic minke whale (*Balaenoptera bonaerensis*) abundance from the IDCR/SOWER data collected from the 1985/86 to the 2003/04 austral summer season. In that study $g(0)$ was estimated less than 1 for Antarctic minke whales (0.327 to 0.793). Using this result, the abundance estimates of Antarctic minke whale based on JARPA data were corrected by application of a regression model (Hakamada *et al.*, 2013).

Preliminary results of the MRDS method on Antarctic minke whales

Data were obtained during the surveys conducted from

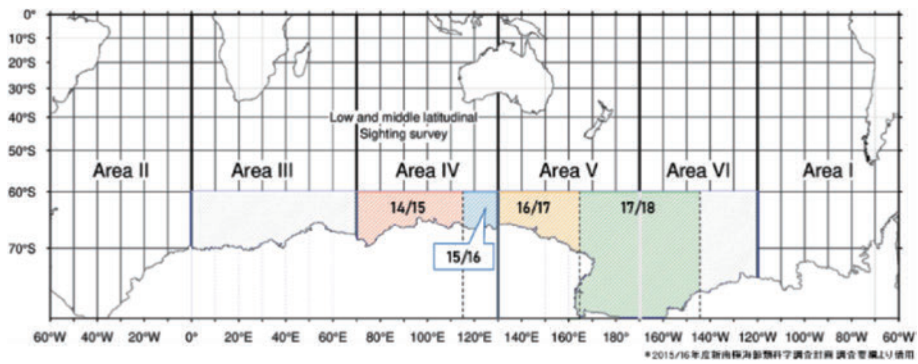

Figure 5. Research area of the study to estimate $g(0)$ of Antarctic minke whale based on MRDS.

December to February from the 2014/15 to the 2017/18 austral summer seasons in Antarctic Area IV, V, and VI (Figure 5). As indicated above, the MRDS method for estimating *g*(0) requires data collected from IO mode survey.

The data comprised a total of 215 schools, including 63 schools which were observed from TOP and IOP (Figure 6, Table 1). Mean school size was 2.23. Also sightings with perpendicular distance of more than 1.5 n.miles, and those observed when Beaufort status was higher than six, were excluded from the dataset.

Table 2 shows the AIC values after fitting explanatory variables to the DS and MR models. For the DS model, school size and cue were the most important explanatory variables. For the MR model, platform and school size were the most important explanatory variables. For the best model, detection probability on the trackline was 0.676 (CV=0.092) for the TOP, 0.429 (CV=0.145) for the IOP, and 0.810 (CV=0.066) for the pooled platforms. The estimated detection function plots are shown in Figure 7.

The detection probability on the trackline for the Antarctic minke whale was similar to the result from the previous study based on the 'OK' method and IDCR-SOWER datasets (Okamura and Kitakado, 2012). In addition, this result was similar to that found for North Pacific common minke whale. For this species the *g*(0) for TOP and upper bridge was estimated at 0.798 by the 'OK' method using the IO passing mode sighting survey data (Okamura *et al.*, 2010).

Future studies

The estimate of *g*(0) for Antarctic minke whale based on the MRDS method was similar to that obtained by the 'OK' method. The estimate considered covariates that affect the detection functions. However, the analysis only accounts for perception bias and did not account for availability bias. Hence, some additional investigations are needed to estimate whale diving pattern as in the case of Heide-Jørgensen *et al.* (2010). They considered availability correction factors using data collected from external experiment survey. Also, model development is needed, for example, correction for non-uniform density of animals within strata using data collected from external experiment survey, and the use of the Bayesian hierarchical approach to incorporate habitat use into a detection model within a mark-recapture distance sampling framework (Oyster *et al.*, 2018), is required. These studies are ongoing at the ICR.

ACKNOWLEDGEMENTS

I thank the researchers, captains, officers and crew members who have participated in all surveys. This work would not have been possible without the substantial efforts of all of them.

Table 1

Number of sightings of Antarctic minke whales made by each platform in each survey. Note that the number of unique sightings is the number of sightings seen by observer 1 plus the number seen by observer 2, minus duplicates.

Survery ID	Number of sightings by observers			Number of unique sightings
	Platform A (TOP)	Platform B (IOP)	Detected by both (Duplicate)	
YS1_1415	14	10	6	18
YS2_1415	6	1	1	6
YS3_1516	47	22	14	55
KY7_1617	6	6	3	9
YS3_1617	14	10	4	20
KY7_1718	10	11	5	16
YS2_1718	79	42	30	91
Total	176	102	63	215

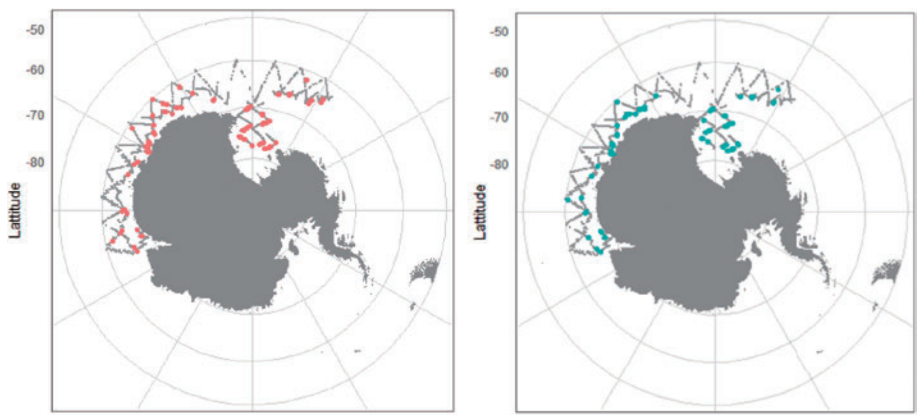

Figure 6. Tracklines and sighting position of Antarctic minke whale observed by each platform (left: TOP; right: IOP). Gray lines indicate tracklines.

Table 2

Akaike Information Criterion (AIC) values after fitting explanatory variables to the Distance Sampling (DS) and Mark-Recapture (MR) models. This table shows a part of all result of model fitted for reference. The final models chosen are given in bold.

Model ID	DS model		MR model	AIC_{DS}	AIC_{MR}	AIC	ΔAIC
	Key*	Covariate	Covariate				
35	hn	beaufort	distance	141.03	472.79	613.82	64.13
37	hn	cue	distance	134.88	472.79	607.67	57.98
31	hn	size	distance	126.90	472.79	599.69	50.01
32	hn	size+beaufort	distance	128.65	472.79	601.44	51.76
33	hn	size+cue	distance	115.75	472.79	588.54	38.85
47	hn	size+cue	distance+platform	115.57	437.56	553.13	3.44
103	hn	size+cue	distance+size	115.57	469.05	584.63	34.94
131	hn	size+cue	distance+beaufort	115.57	473.72	589.30	39.61
145	hn	size+cue	distance+cue	115.57	474.53	590.11	40.42
54	**hn**	**size+cue**	**distance+platform+size**	**115.57**	**434.11**	**549.69**	**0.00**
61	hn	size+cue	distance+platform+beaufort	115.57	438.48	554.06	4.37
68	hn	size+cue	distance+platform+cue	115.57	439.53	555.10	5.42
40	hr	size+cue	distance	118.10	472.79	590.89	41.21
159	hr	size+cue	distance+platform+size	118.10	434.11	552.21	2.53

*: hn denotes a Half-normal function; hr denotes a Hazard-rate function.

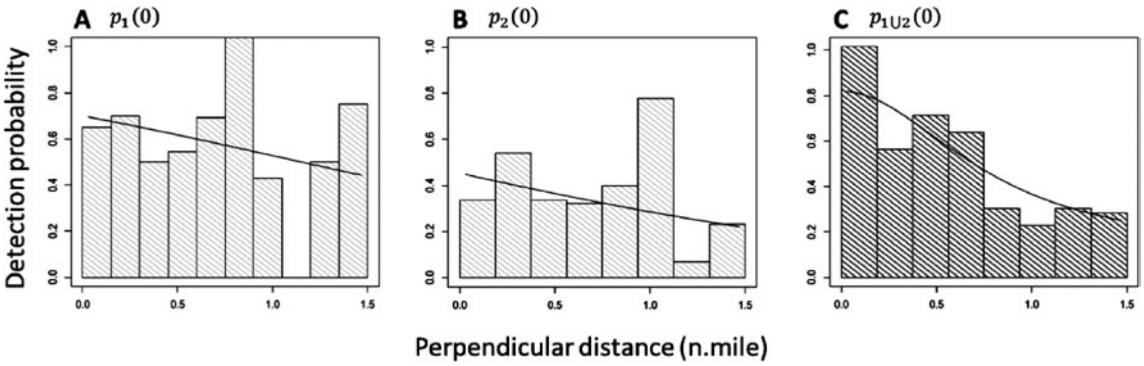

Figure 7. Detection function plots for the TOP (A), IOP (B) and both platforms pooled (C) by the MRDS method.

REFERENCES

Akaike, H. 1973. Information theory as an extension of the maximum likelihood principle. pp. 267–281. *In*: B.N. Petrov and F. Csaki (eds.) *Second International Symposium on Information Theory*. Akademiai Kiado, Budapest. 451 pp.

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.

Barlow, J. 1999. Trackline detection probability for long-diving whales. pp. 209–221. *In*: G.W. Garner, S.C. Amstrup, J. Laake, G.F.J. Manly, L.L. McDonald and D.G. Robertson (eds.) *Marine Mammal Survey and Assessment Methods*. Balkema, Rotterdam, the Netherlands, 287 pp.

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

Borchers, D.L., Laake, J.L., Southwell, C. and Paxton, C.G.M. 2006. Accommodating unmodeled heterogeneity in double-observer distance sampling surveys. *Biometrics* 62: 372–378.

Buckland, S.T. and Turnock, B.J. 1992. A robust line transect method. *Biometrics* 48: 901–909.

Buckland, S.T., Anderson, D.R., Burnham, K.P. and Laake, J.L. 1993. *Distance Sampling*. London: Chapman and Hall, reprinted 1999 by RUWPA, University of St Andrews. 446 pp.

Buckland, S.T., Anderson, D.R., Burnham, K.P., Laake, J.L., Borchers, D.L. and Thomas, L. 2001. *Introduction to Distance Sampling: Estimating Abundance of Biological Populations.* Oxford: Oxford University Press. 432 pp.

Burt, M.L., Borchers, D.L., Jenkins, K.J. and Marques, T.A. 2014. Using mark-recapture distance sampling methods on line transect surveys. *Methods in Ecology and Evolution* 5: 1180–1191.

Butterworth, D.S. and Borchers, D.L. 1988. Estimates of *g*(0) for minke schools from the results of the independent observer experiment on the 1985/86 and 1986/87 IWC/IDCR Antarctic assessment cruises. *Rep. int. Whal. Commn* 38: 301–313.

Cooke, J.G. 2001. A modification of the radial distance method for dual-platform line transect analysis, to improve robustness. Paper SC/53/IA31 presented to the IWC Scientific Committee, July 2001 (unpublished). 7 pp. [Available from the IWC Secretariat].

Dawson, S.M., Slooten, E., DuFresne, S.D., Wade, P.R. and Clement, D.M. 2004. Small-boat surveys for coastal dolphins: line-transect surveys of Hector's dolphins (*Cephalorhynchus hectori*). *Fishery Bulletin* 102: 441–451.

Hakamada, T. and Matsuoka, K. 2017. Sighting survey procedures for abundance estimates of large whales in JARPA and JARPAII, and results for Antarctic minke whales. *Technical Report of the Institute of Cetacean Research (TEREP-ICR)* No. 1: 28–36.

Hakamada, T., Matsuoka, K., Nishiwaki, S. and Kitakado, T. 2013. Abundance estimates and trends for Antarctic minke whales (*Balaenoptera bonaerensis*) in Antarctic Area IV and V for the period 1989/90–2004/05. *J. Cetacean Res. Manage.* 13: 123–151.

Heide-Jørgensen, M.P., Laidre, K.L., Simon, M., Burt, M.L., Borchers, D.L. and Rasmussen, M. 2010. Abundance of fin whales in West Greenland in 2007. *J. Cetacean Res. Manage.* 11: 83–88.

Innes, S., Heide-Jørgensen, M.P., Laake, J.L., Laidre, K.L., Cleator, H., Richard, P. and Stewart, R.E.A. 2002. Surveys of belugas and narwhals in the Canadian high Arctic in 1996. *Scientific Publications of the North Atlantic Marine Mammal Commission* 4: 169–190.

International Whaling Commission. 2012. Requirements and Guidelines for Conducting Surveys and Analysing Data within the Revised Management Scheme. *J. Cetacean Res. Manage.* (Suppl.) 13: 509–517.

Laake, J.L. and Borchers, D.L. 2004. Methods for incomplete detection at distance zero. pp. 108–189 *In*: S.T. Buckland, D.R. Anderson, K.P. Burnham, J.L. Laake, D.L. Borchers and L. Thomas (eds.) *Advanced Distance Sampling*. Oxford University Press, Oxford. 416 pp.

Matsuoka, K., Ensor, P., Hakamada, T., Shimada, H., Nishiwaki, S., Kasamatsu, F. and Kato, H. 2003. Overview of minke whale sightings surveys conducted on IWC/IDCR and SOWER Antarctic cruise from 1978/79 to 2000/01. *J. Cetacean Res. Manage*. 5 (2): 173–201.

Marsh, H. and Sinclair, D.F. 1989. Correcting for visibility bias in strip transect aerial surveys of aquatic fauna. *J. Wildlife Manage.* 53: 1017–1024.

Okamura, H. and Kitakado, T. 2012. Abundance estimates of Antarctic minke whales using the OK method. Paper SC/64/IA2 presented to the IWC Scientific Committee, June 2009 (unpublished). 24 pp. [Available from the IWC Secretariat].

Okamura, H., Kitakado, T., Hiramatsu, K. and Mori, M. 2003. Abundance estimation of diving animals by the double-platform line transect method. *Biometrics* 59: 512–520.

Okamura, H., Miyashita, T. and Kitakado, T. 2010. *g*(0) estimates for western North Pacific common minke whales. Paper SC/62/NPM9 presented to the IWC Scientific Committee, June 2010 (unpublished). 7 pp. [Available from the IWC Secretariat].

Oyster, J.H., Keren, I.N., Hansen, S.J.K. and Harris, R.B. 2018. Hierarchical Mark-Recapture Distance Sampling to Estimate Moose Abundance. *J. Wildlife Manage.* 82: 1168–1679.

Palka, D. 1995. Abundance estimate of the Gulf of Maine harbor porpoise. *Rep. int. Whal. Commn* (special issue) 16: 27–50.

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* 1: 170–189.

Schweder, T., Skaug, H.J., Dimakos, X.K., Langaas, M. and Oien, N. 1997. Abundance of northeastern Atlantic minke whales, estimates for 1989 and 1995. *Rep. int. Whal. Commn* 47: 453–483.

Skaug, H.J. and Schweder, T. 1999. Hazard models for line transect surveys with independent observers. *Biometrics* 55: 29–36.