Proceedings of the workshop on

Estimation of g(0) in line-transect surveys of cetaceans

Held at the European Cetacean Society's 18th Annual Conference,
Vildmarkshotellet at Kolmården DjurPark, Kolmården, Sweden,
28th March 2004

Editors:

Frank Thomsen[1,2], Fernando Ugarte[3,4] and Peter G.H. Evans[4,5]

[1] Biologisch-landschaftsökologische Arbeitsgemeinschaft (biola), Gotenstraße 4, D-20097 Hamburg, Germany
[2] Biozentrum Grindel, Universität Hamburg, Martin-Luther-King-Platz 3, D-20146 Hamburg, Germany
[3]Cardigan Bay Marine Wildlife Centre, New Quay, Ceredigion SA45 9PS Wales, UK
[4]Sea Watch Foundation, 11 Jersey Road, Oxford 0X4 4RT, UK
[5]Dept. of Zoology, University of Oxford, South Parks Road, Oxford 0X1 3PS, UK

ECS NEWSLETTER NO. 44 – SPECIAL ISSUE
April 2005

SHIPBOARD SURVEYS IN THE NORTHWEST ATLANTIC: ESTIMATION OF *g(0)*

D. Palka

National Marine Fisheries Service, 166 Water St., Woods Hole, MA USA 02543
(e-mail: debra.palka@noaa.gov)

INTRODUCTION

Since 1991, staff from the Northeast Fisheries Science Center of the U.S. National Marine Fisheries Service have been conducting shipboard cetacean abundance surveys that provide an estimate *g(0)*, the probability of detecting a group on the track line. To estimate *g(0)*, two teams on the same ship simultaneously collect standard line-transect data. These data are then analysed using the direct duplicate method (Palka, 1995), a type of sight-resight (mark-recapture) analysis.

During observer sight-resight surveys, animals can be missed at any distance from an observer. If heterogeneity exists in the probability of detecting animals, the resulting abundance estimates will be biased. To account for such heterogeneity, the direct duplicate method presented in this paper assumes that detections are independently conditional on perpendicular distance (x) and other covariates. Typically, only perpendicular distance is used to explain heterogeneity in the detection probability. However, when important covariates are excluded, the remaining heterogeneity generates biased abundance estimates. The direct duplicate estimator, based on the standard Lincoln-Petersen sight-resight estimator, assumes that detections on the track line (x=0) made by multiple observers are independent; this is referred to as "track line conditional independence" (Laake, 1999). As such, covariates, other than perpendicular distance, are only needed to account for heterogeneity in the detection of groups on the track line. The track line conditional independence assumption is less restrictive than full conditional independence, which presumes no heterogeneity at any distance from the track line. This latter assumption is implicit in some other abundance estimation methods, such as those by Manly *et al.* (1996) and Borchers *et al.* (1998).

In this paper, I describe the direct duplicate method and show how covariates can be incorporated into the analysis. I then provide estimates of *g(0)* for cetaceans in pelagic waters off the east coast of the U.S., based on data collected during a 1998 cetacean abundance line-transect sighting survey.

MATERIALS AND METHODS

Field Data Collection Methods Data were collected by two "independent" sighting teams aboard the ship *R/V Abel-J*, which travelled at 17-19 km/hr (9-10 knots) during survey operations. Surveying was conducted when Beaufort sea states were four and less and was continuously performed, whenever weather conditions allowed, between 6 am and 6 pm, with one hour off for lunch. Observers searched the waters from directly in front of the vessel to 90 degrees left and right of the track line, and from the ship out to the horizon. Each team consisted of four people: three on-duty and one at rest. The upper team was located in a crow's nest 14 m above the sea surface and the lower team was located on top of the bridge, 8.5 m above the sea surface. To ensure animals were seen before they reacted to the ship, two of the three on-duty observers used binoculars to search far from the ship while the third observer recorded data and scanned with naked eye closer to the ship. Because of the physical size of the sighting platforms, the upper team searched with 20x60 binoculars and the lower team searched with 25x150 binoculars.

Sightings data were recorded on hand-held data entry computers (e.g., Garrett-Logan and Smith, 1997). For each sighting, the following data were recorded: time of initial detection, bearing and distance to the group, species composition of the group, best, high and low estimate of group size, group behaviour, sighting cue, and swim direction. To determine which groups were detected by both the upper and lower teams (duplicate sightings), data on time, position, and swim direction were recorded for subsequent surfacings of as many groups as possible.

A computer on the bridge connected to a differential GPS and bridge instruments recorded the following once per minute: time, wind speed and direction, depth, surface temperature, surface drift speed and direction, and the ship's position, speed and course. In addition, the following were recorded every half hour or whenever conditions changed: time, observer's position, swell height and direction, Beaufort sea state, direction of sun, magnitude of glare, and visibility.

Direct Duplicate Analysis Method

Using the direct duplicate estimator, density, corrected for *g(0),* for species *i*, D_{ci}, is calculated as:

$$D_{ci} = \frac{D_{i.upper} \cdot D_{i.lower}}{D_{i.duplicate}} \tag{1}$$

where $D_{i.upper}$ is the density of species *i* using only the upper team's data, $D_{i.lower}$ is the density using only the lower team's data, and $D_{i.duplicate}$ is the density using only data from sightings seen by both teams (duplicate sightings). Each of these three densities are estimated in the usual way:

$$D = \frac{n \cdot E(s)}{2L \cdot ESHW} \tag{2}$$

For example, density of duplicate sightings, $D_{i.duplicate}$ is where *n* is the number of detected duplicate groups, *E(s)* is the expected group size of duplicate sightings, *L* is the length of the track line and *ESHW* is the estimated strip half width for duplicate sightings. Thus, *g(0)* for species *i* as seen from the upper team, $g(0)_{i.upper}$, and lower team, $g(0)_{i.lower}$, are derived as follows:

$$g(0)_{i.upper} = \frac{n_{i.duplicate} \cdot ESHW_{i.lower}}{n_{i.lower} \cdot ESHW_{i.duplicate}} \quad and \quad g(0)_{i.lower} = \frac{n_{i.duplicate} \cdot ESHW_{i.upper}}{n_{i.upper} \cdot ESHW_{i.duplicate}} \tag{3}$$

After the sightings data were collected, duplicate sightings were determined using a Visual Basic program that, for the time of each sighting, mapped the position of the sighting relative to the ship and predicted positions of previous sightings from both teams. Predicted positions were calculated using swim direction, time, previous sighting locations, ship's speed, and estimated animal swim speed. Swim speeds were adjusted for different species. Coefficients of variation (CV) of density and *g(0)* estimates were derived using bootstrap re-sampling techniques (Efron and Tibshirani, 1993).

Incorporating Covariates into the Direct Duplicate Estimator

Covariates can easily be incorporated into the above formulas by inclusion within any of the density estimates in Equation 1. A different set of covariates can be used for each team and for duplicate sightings. The covariate models in DISTANCE 4 (Buckland *et al.,* 2001) were used in the present study to estimate the three densities (*i.e.,* the appropriate *ESHW*'s in Equation 2). The minimum AIC criterion was used to select the best set of covariates.

Choices of covariates included the following animal-related metrics (group size, group behavior and initial cue) and survey-related metrics (observer, Beaufort sea state, wind speed and water depth). Group size was defined as the best estimate of group size. Group behaviour was defined as the behaviour when the group was initially detected and was originally recorded in one of 12 categories. For this analysis, behaviour categories were lumped into three different levels of activity: swimming (low profile types of activities including swimming, feeding, logging and milling), porpoising (medium profile activities including porpoising, bow riding, courting, diving and fluking), and charging (high profile activities including charging, aerobatics and breaching). Initial cue was defined as the behavioural phenomenon that caught the attention of the observer when the group was initially detected (body, splash or blow). Observer was defined as the specific individual who initially detected the group. Because wind speed and Beaufort sea state are highly correlated, the detection function model was not allowed to include both. Depth was water depth (in metres), as measured at the ship's location when the group was initially detected.

RESULTS

During July 6 to August 4, and August 8 to September 6, 1998, the *R/V Abel-J* surveyed: (a) continental shelf edge waters between the 100 m and 2000 m isobaths, and (b) the Gulf Stream and off-shelf waters greater than the 2000 m depth contour (Fig. 1). The ship surveyed 2,985 km in the shelf edge stratum (area = 55,798 km^2) and 1,429 km in the off-shelf stratum (area = 113,201 km^2). Sixteen species of whales, dolphins, and porpoises were detected. Species included in this analysis were those where more than ten groups were detected by a team: common dolphins (*Delphinus delphis*), Risso's dolphins (*Grampus griseus*), offshore bottlenose dolphins (*Tursiops truncatus*), striped dolphins (*S. coeruleoalba*), fin whales (*Balaenoptera physalus*), sperm whales (*Physeter macrocephalus*), pilot whale spp. (*Globicephala* spp.), and beaked whales (*Mesoplodon* spp. and *Ziphius cavirostris*).

For all species except Risso's dolphins, estimates of *g(0)* were generally higher for the upper team than for the lower team (Table 1). Without covariates, *g(0)* estimates ranged from 0.28 for sperm whales to 0.99 for offshore bottlenose dolphins. At least one covariate was found significant for each species, except for Risso's dolphins. Beaufort sea state (or wind speed) was the most common significant covariate, followed by group size and initial cue. When covariates were included, all estimates of *g(0)* decreased or remained the same (to within two decimal places), except for the pilot whale estimates which increased. A decrease in *g(0)* implies an increase in the abundance estimate. Excluding pilot whales, on average there was a 12% reduction in the *g(0)* estimate with covariates, versus without them.

It is presumed that the estimate of *g(0)* for pilot whales increased when covariates were included because this species was attracted to the ship. Attraction is a common pilot whale behaviour; and, when accounted for, will reduce the abundance estimate. And indeed, this is what happened when covariates were incorporated into the *g(0)* estimate.

DISCUSSION AND CONCLUSIONS

In this study, the effect of excluding covariates resulted in a negative bias of about 12% in the estimate of species abundance. This pattern has been shown in other studies. For example, Schweder (1999) demonstrated that without accounting for heterogeneities in weather, sea state, and observer skill, the estimated abundance of Northeast Atlantic minke whales was negatively biased by 27%.

Other potential covariates that might be considered are habitat metrics, such as plankton types and densities, salinity, bottom slope, and distance from a front. The covariate ‘observer’ was not found to be significant in the present study presumably because there were too many observers and thus too many degrees of freedom to detect differences large enough to be significant. Perhaps a more efficient way to incorporate observer effects (if they exist) is to create, say, three categories of observers: highly experienced, intermediate level of experience, and little experience. Experience could be measured as the amount of previous time spent conducting abundance surveys, or as a function of the observer’s sighting rate.

Responsive movements and dive patterns should also be accounted for to derive the least biased estimates of *g(0)* and density (or abundance). Given the responsive attraction of pilot whales towards the survey ship, a more appropriate method to estimate *g(0)* and density for this species is the Buckland-Turnock analysis method (Buckland and Turnock, 1992). There are two types of bias that could occur when an animal is missed. Perception bias occurs when an animal is missed because the observer simply did not recognise it, even though it was at the surface. Availability bias occurs when an animal is missed because it was submerged during the entire time that the ship was passing by. The direct duplicate method described in this paper can account for perception bias, but not availability bias. For example, long diving animals, such as sperm whales and beaked whales, can be submerged for up to an hour, and when this happens there is generally no opportunity for either team to make a sighting. To derive a less biased abundance estimate for these species, it is necessary to include dive time patterns into the estimation of *g(0)*, as was done, for example, by Schweder *et al.* (1999).

REFERENCES

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(4)**: 1221-1237.

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., Laake, J.L., Borchers, D.L. and Thomas, L. 2001. *Introduction to Distance Sampling*. Oxford University Press, Oxford. 432pp.

Efron, B. and Tibshirani, R.J. 1993. *An Introduction to the Bootstrap*. Chapman and Hall. New York. 436pp.

Garrett-Logan, N. and Smith, T. 1997. A hand-held, pen-based computer system for marine mammal sighting surveys. *Marine Mammal Science* **13(4)**: 694-700.

Laake, J. 1999. Distance sampling with independent observers: Reducing bias from heterogeneity by weakening the conditional independence assumption. Pp. 137-148. In: *Marine Mammal Survey and Assessment Methods* (Eds. Garner, G.W., S.C. Amstrup, J.L. Laake, B.F.J. Manly, L.L. McDonald and D.G. Robertson) A.A. Balkema. 287pp.

Manly, B.F.J., McDonald, L.L. and Garner, G.W. 1996. Maximum likelihood estimation for the double-count method with independent observers. *J. of Agri, Biol and Envir Stat* **1**: 170-189.

Palka, D. 1995. Abundance estimate of the Gulf of Maine harbour porpoise. Pp. 27-50. In: *Biology of the Phocoenids, Special issue 16* (Eds. Bjørge, A. and Donovan, G.P.).

Schweder, T., Skaug, H.T., Langaas, M. and Dimakos, X. 1999. Simulated likelihood methods for complex double-platform line transect surveys. *Biometrics* **55(3)**: 120-128.

Table 1. For each species or species group, estimates of *g(0)*, with and without covariates (covs), and covariates that were significant for the upper and lower teams, and for duplicate sightings. Covariate abbreviations are: Beaufort = Beaufort sea state, Wind = wind speed, Size = group size, Behaviour = behaviour category, Cue = cue category, Temp = water temperature, and None = no significant covariate

Species	Team	Number of groups	*g(0)*		Significant covariates
			Without covs	With covs	
Beaked whales	Upper	53	0.50 (0.41)	0.50 (0.66)	Beaufort
	Lower	58	0.51 (0.40)	0.46 (0.63)	None
	Duplicates	17	-	-	None
Bottlenose dolphin	Upper	61	0.99 (0.84)	0.93 (0.61)	None
	Lower	79	0.69 (0.80)	0.69 (0.58)	Temp, Wind
	Duplicates	36	-	-	None
Common dolphin	Upper	30	0.95 (0.30)	0.52 (0.89)	Cue
	Lower	42	0.87 (0.39)	0.76 (0.87)	Cue, Wind
	Duplicates	20	-	-	None
Fin whales*	Upper	27	0.80 (0.59)	0.68 (0.80)	Beaufort, Cue
	Lower	10	0.44 (0.47)	0.32 (0.87)	Size, Wind
	Duplicates	7	-	-	None
Risso's dolphin	Upper	122	0.51 (0.36)	0.51 (0.36)	None
	Lower	144	0.61 (0.31)	0.61 (0.31)	None
	Duplicates	58	-	-	None
Sperm whale	Upper	63	0.42 (0.34)	0.36 (0.66)	None
	Lower	49	0.28 (0.34)	0.28 (0.67)	Cue
	Duplicates	21	-	-	None
Striped dolphin	Upper	45	0.77 (0.29)	0.76 (0.77)	Wind
	Lower	53	0.68 (0.32)	0.61 (0.77)	Size
	Duplicates	32	-	-	Size
Pilot whales	Upper	36	0.52 (0.86)	0.59 (0.68)	Behaviour
	Lower	40	0.48 (0.55)	0.50 (0.65)	None
	Duplicates	20	-	-	Cue

* *ESHW* was estimated from data pooled over groups identified as fin whale, fin or sei whale (*Balaenoptera borealis*) (7 groups detected by the upper and 7 groups detected by the lower team) and sei whale (2 groups detected by upper team).

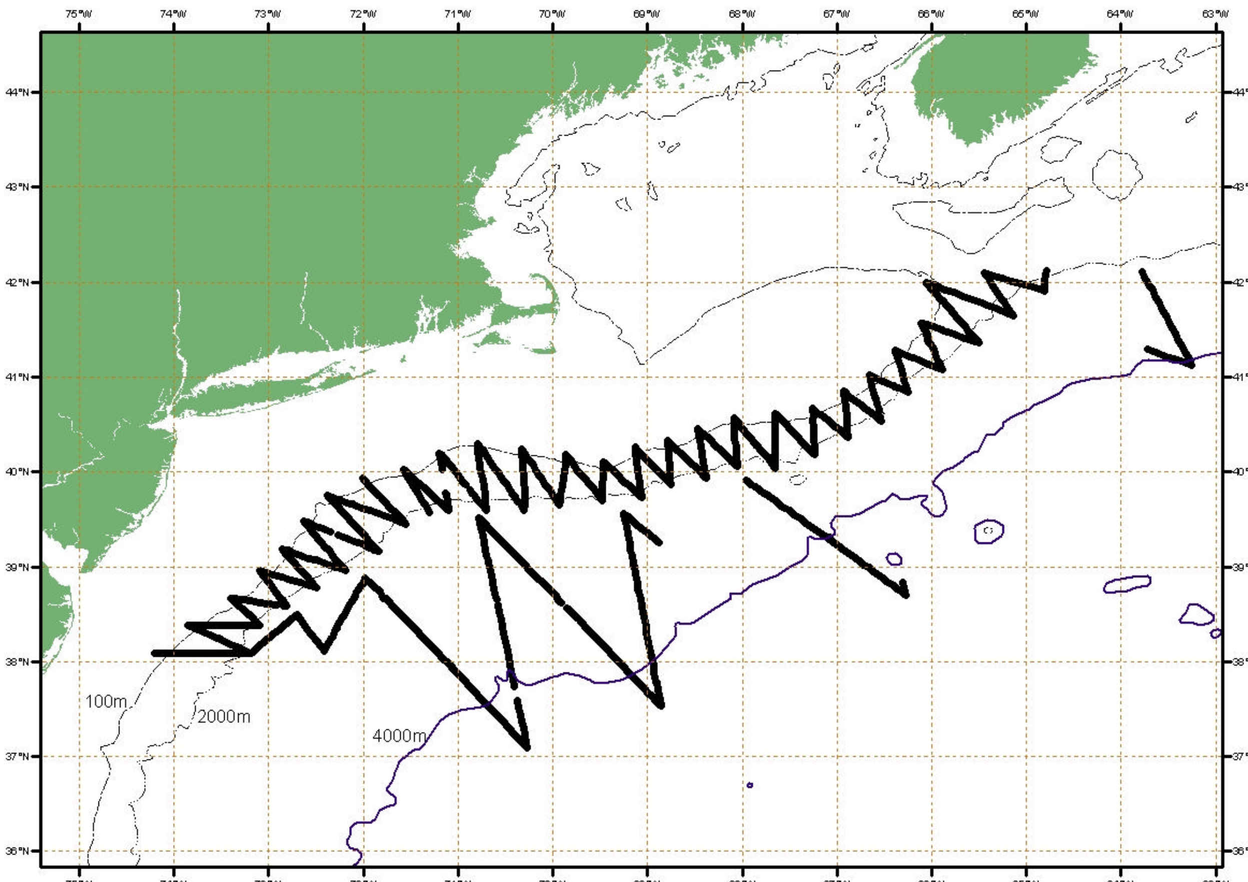

Fig. 1. Track lines surveyed by the *R/V Abel-J* during the 1998 cetacean abundance survey