Progress on data analysis from a 2002 Amazon river dolphin survey for boto and tucuxi in Colombia and Peru

SC/64/SM24

Rob Williams, Louise Burt, Sharon Hedley, Fernanda Marques, Fernando Trujillo

rmcw@st-andrews.ac.uk

Introduction and rationale for IWC SC meeting document

A survey of Amazon river dolphins (*Inia geoffrensis* and *Sotalia fluviatilis*) was conducted in March/April, 2002, on a section of the River Amazon which forms a border between Columbia and Peru. A combination of strip transect and line transect methodology was used. There were two observation platforms; one at the front and one in the rear where observers looked behind the boat. This observer configuration allowed the probability of detection of the front platform to be estimated and so density could be estimated without assuming that detection on the trackline was certain.

At its 2000 meeting, the Small Cetaceans subcommittee "noted that few reliable estimates of abundance were available for any species of freshwater cetacean and that the habitat and behaviour of these species posed particular problems for abundance estimation" (IWC 2001, JCRM 3 (Suppl.), p. 277). The subcommittee recommended that "scientists with appropriate theoretical and/or analytical skills should be directly involved in river cetacean studies, so that surveys result in statistically robust estimates of abundance" (ibid.). In 2002, two biologists and two statisticians led a pilot survey of boto (*Inia geoffrensis*) and tucuxi (*Sotalia fluviatilis*) in Colombian and Peruvian stretches of the Amazon, which was reported to the subcommittee in 2003 (IWC 2003, JCRM 5 (Suppl.), p. 371). At that time, no analyses had taken place, but the "sub-committee thanked Hedley and Williams for their efforts and agreed that these activities will provide useful information to groups studying freshwater cetaceans and other species inhabiitng complex environments, such as *Sousa*. The sub-committee recommended the continued development of these techniques."

We were inspired to revisit this dataset in light of reports of river dolphins being caught for use as fishery bait, not only in Brazil (da Silva & Martin 2000), but also in our study area (Gomez-Salazar et al. 2012).

A survey of Amazon river dolphins (*Inia geoffrensis* and *Sotalia fluviatilis*) was conducted from 28 March to 5 April, 2002, along the section of the River Amazon that forms a border between Colombia and Peru, from Atacuari to Leticia (Figure 1). The survey area also included a few tributaries and lakes, although as there were only nine days available for the survey, it was not possible to dedicate much time to tributaries. The timing of the survey was chosen to be in the high water season – the rationale being that this facilitated access to most sections of the river and associated waters.

Prior to this survey, a pilot survey was carried out in April 2001, on a similar vessel in the same survey area. The survey area also coincides substantially with that of (Vidal et al. 1997), who present the only previous estimates of abundance of the two species in this area, although their study extended a few kilometres beyond Leticia into Brazilian waters, and was conducted during the low water season in June.

The main objectives of the survey were twofold:

1. to obtain abundance estimates of *I. geoffrensis* and *S. fluviatilis* in the study area.
2. to develop protocols and make recommendations for future surveys of river dolphins, particularly when budgetary and other resource limitations apply.

This cruise report is intended outline some of the problems faced when surveying river dolphins, to document the methods used in this survey, and to briefly summarize some initial thoughts and recommendations for repeat, or other, river dolphin surveys. The data themselves are currently undergoing validation screening, therefore any results presented here should be considered preliminary. Detailed analyses will appear elsewhere.

Methods

Field methods

The survey vessel was a 22.2m river boat, the *Principe Edifa*, based in Tabatinga, Brazil. Her draft was just 0.7m, allowing navigation in shallow waters. She was powered by two 80 Hp diesel engines, which was usually sufficient to maintain speeds of at least 7 km/h, even against strong currents. Nevertheless, speed did vary considerably according to current direction and strength, but was typically between 7 and 13 km/h.

Two observation platforms were constructed. The front (forward-facing) platform was positioned on the roof of the vessel, above the wheelhouse, with two seated primary observers searching with naked eye (at eye height of approximately 6.4m), one central standing observer who searched with naked eye and with 7x binoculars, and a data recorder. The stern (rear-facing) platform was positioned on the middle level of the vessel, with two seated observers searching with naked eye, at eye height of approximately 4.4m.

SURVEY METHODS

The river was divided into six different strata; river, tributary, channel (narrow section of the main river), lake, island and confluence (within 150m of a junction of two water types). Transects tended to be aligned parallel to the bank unless the river was wide enough to allow transects to lie perpendicular to bank (eg. in lakes).

The location and environmental conditions were recorded approximately every 10 minutes of search effort. On detecting a group, the radial distance, sighting angle, group size and species were recorded. The sighting angles and radial distance for the front platform were corrected for observer bias (Williams *et al*. 2007). The sightings made by the rear platform have not been corrected. Duplicate sightings, those seen by both the front and rear platforms, were assessed by timing and location, such that the two platforms were not completely independent. Instead, the observers on the two platforms communicated by 2-way radio to assess certainty of a duplicate detection.

ANALYSIS METHODS

Classifying line type

The width of the river varied and where the river was narrow or the vessel was close to one riverbank, the observers felt that there was a constant level of detection across the river or to the riverbank (ie. strip transect sampling). For this analysis, if the distance from the vessel to the river bank was 100m or less, then the transect was defined to be a strip transect. If the vessel was more than 100m from the riverbank, then it was assumed that not all animals were being detected and so the transect was defined to be a line transect (Buckland *et al*. 2001) was used. Frequently this meant that a strip transect operated on one side of the vessel with a line transect on the other side (Table 1).

Line transect sampling estimators

Group density for each line type (D_s)was estimated as follows:

$$\hat{D}_s = \frac{n}{2wL\hat{p}} \quad (1)$$

where n is the number of groups detected in the stratum and $\hat{p}$ is the estimated probability of detecting a group (see below). Individual animal density (D) was obtained from

$$\hat{D} = \hat{D}_s \bar{s}$$

where $\bar{s}$ is mean group size.

The variance of the encounter rate (n/L) was estimated using the method developed by Innes *et al*. (2002) using the R2 form of the estimator as in Fewster *et al*. (2009) - the default estimator in Distance (Thomas *et al*. 2010).

Probability of detection

Having two teams of observers (front and rear) allowed a mark-recapture distance sampling (MRDS) approach to be used to estimate the probability of detection (Laake and Borchers 2004). An MRDS approach requires two subsidiary models to be fitted; a distance sampling (DS) model obtained from the perpendicular distance distribution assuming that detection on the trackline is certain (ie $g(0)=1$) and a mark-recapture (MR) model to obtain the probability of detection on the trackline. This probability of detection at perpendicular distance zero is then used to adjust the DS detection function to obtain an overall probability of detection. The observing teams in this survey acted independently, however, in this case, it was felt that there was likely to have been responsive movement of the animals by the time the animals were detected by the rear platform and so the DS model was fitted to sightings made by the front platform only. Thus, an estimate of the probability of detection on the trackline for the front platform was required from the MR model.

For the DS model, both a hazard-rate ($1-\exp(-x/\sigma)^{-b}$) and a half-normal form ($\exp(-x^2/2\sigma^2)$) were considered as suitable forms for the detection function (where σ is a scale parameter, x is perpendicular distance and b is a shape parameter) (Buckland *et al*. 2001). Akaike's Information Criterion (AIC) and goodness of fit statistics were used to select the final model and all model selection was performed in the program Distance (Thomas *et al*. 2010; version 6.1 Beta 1 and version 2.0.6 of the mrds R library).

The MR detection function defines the probability that an animal at perpendicular distance x and covariates z was detected by the front platform, given that it was seen by the rear, and is denoted by $p_{1|2}(x,z)$ and is generally modelled using the logistic form:

$$p_{1|2}(x,z) = \frac{\exp\left\{\beta_0 + \beta_1 x + \sum_{k=1}^{K} \beta_{k+1} z_k\right\}}{1+\exp\left\{\beta_0 + \beta_1 x + \sum_{k=1}^{K} \beta_{k+1} z_k\right\}} \quad (2)$$

where β_0, β_1, , β_{K+1} represent the parameters to be estimated and K is the number of covariates other than distance. In a typical trial configuration set up, it is assumed that platform setting up the trials detects the animals before they have responded to the vessel. In this case, the sightings from the rear platform were used as the trials (ie did the front platform detect the same animal or group or not) and because distribution of animals was unlikely to have been uniform (with respect to the trackline) by the time the rear platform detected an animal, perpendicular distance was ignored and not included as a covariate in the above model. Note that in the simplest case with no covariates, this will revert to the Petersen estimator.

Species and group size were potential explanatory variables for inclusion into both the DS and MR models and Akaike's Information Criterion (AIC) and goodness of fit tests were used for model selection.

Strip transect estimator

In strip transect methodology, it is assumed that all sightings within the strip are detected. In this survey the probability of detection by the front platform can be estimated using the rear platform observations as before in equation 2 and so the term 'strip transect' is used here to mean that the probability of detection is constant out to a distance of 100m but is not necessarily certain. Thus for the strip transects (or half transects), then density was estimated as in equation 1 with the probability of detection, $\hat{p}$, including only a component for $p_{1|2}(z)$ (again excluding perpendicular distance as a covariate).

Combining strip/line transects

Group density in a stratum was (D_s) was obtained from

$$\widehat{D}_s = \frac{{n_s}/{\hat{p}_s} + {n_l}/{\hat{p}_l} + {n_{ts}}/{\hat{p}_s} + {\hat{n}_{tl}}/{\hat{p}_l}}{2w_sL_s + 2w_lL_l + L_t(w_s + w_l)}$$

where the subscripts *s* and *l* refer to strip and line, respectively, and *ts* refers to a sightings from the strip side of a strip/line transect and *tl* refers to a sightings from the line transect side of a strip/line transect.

Note another approach would be to calculate overall density in a stratum by taking a weighted average of the densities for each line type, where the lengths of the line types were used as the weights (possibly more difficult to disentangle terms to obtain CV?)

Density gradient of animals

In conventional distance sampling (Buckland *et al*. 2001), it is generally assumed that, given random line placement, that animal groups are uniformly distributed with respect to the perpendicular distance. If dolphins tended to prefer the region in the middle of the river or the river banks, then this assumption may be violated with transects running parallel to the riverbank. Also, imperfect detection may be confounded with a density gradient across the river (ie. detection may decrease away from the vessel but if density also decreases, then the decrease in detection may also be due to a decrease in the number of animals). Distances to the port and starboard shores were routinely recorded during the survey and so the width of the river could be calculated. Note that the range finder could only measure accurately to 500m and so all distances recorded as being >500m were set to 500m. Given the perpendicular distance of the animal and the side of the vessel the animal was recorded on, the distance of the sighting to the port bank could be calculated. Since the river width changes, then the distribution of groups across the river (from the port bank) was calculated as follows

$$proportion\ across\ river = \frac{distance\ from\ port\ bank}{river\ width}$$

Thus, for an animal close to the port bank, the proportion across the river will be close to zero and for an animal closer to the starboard bank, the proportion will be closer to one.

RESULTS

If the distance to the port or starboard river bank had not been recorded, then the previous non-missing value on the same transect was used. The majority of search effort was conducted with a strip transect operating on one side of the vessel and a line transect on the other (Table 1 and Figure 1). A summary of the number of sightings is shown in Table 2. Six detections seen by the front observers have been excluded from analyses as no perpendicular distances were obtained (including all groups detected in a confluence and so this stratum is excluded from further analysis). Two detections by the rear observers have been excluded because they were seen off effort.

The histograms of perpendicular distance for each platform are shown in Figure 3.

Probability of detection

Line transect sightings

Sightings where the species was undetermined have been excluded and the perpendicular distances have been truncated at 200m. The fitted MR model is shown in Figure 4 and both species and group size were selected as explanatory variables. (Out of interest, perpendicular distance was also included as an exploratory variable but was not selected by AIC.) The probability of detection on the trackline for the front platform, given that it was seen by the rear, was 0.66 (CV=0.08). For the DS model a hazard rate form was selected with no additional explanatory variables. Combining the intercept from the MR model and the shape from the DS model, resulted in a probability of detection for the front platform of 0.34 (0.15) (Figure 5).

Strip transects

For sightings detected during strip transects, no explanatory variables were chosen in the model (so this was equivalent to fitting a Petersen estimator) and the probability of detection by the front platform was 0.81 (0.11).

Density estimates

Density estimates for each stratum are given in Table 5. A mean group size, pooling all sightings of Sotalia and Inia within 200m, was 1.76 (CV=0.05) and used to convert group density to animal density.

Density gradient across the river

Using all front sightings (where distances to both banks were recorded), the distribution of groups across the river is shown in Figure 6. The sample size is not large but it appears that groups prefer a region close, but not close, to the river bank.

DISCUSSION

Density estimates were obtained for river dolphins in five habitat types. Both species were combined and a single estimate of group size was obtained. Table 3 indicates that group size may be different for each species with *Sotalia* having a slightly larger group size than *Inia*.

More work on density gradient required.

Low power to detect declines is a recurring theme in marine mammal conservation biology (Taylor et al. 2007), but traditional survey methods appear to be particularly poor at detecting declines in freshwater cetacean populations (Huang et al. 2012). In light of that difficulty, specialists in this taxon encourage regulatory agencies to adopt precautionary conservation measures at the level of the population, ecosystem and habitat long before signs of population and habitat loss are apparent, in order to develop risk-averse conservation measures for freshwater cetaceans (Huang et al. 2012) (Gomez-Salazar et al. 2012).

Acknowledgements

We would like to thank Andy Read for initially suggesting this collaboration between river dolphin biologists and statisticians. We thank Maria-Claudia Diazgranados, Tom Hedley & Leo Vicari for field data collection; David Borchers for advice on analysis of double-platform sightings data; Victoria Mars for funding for field work & Sculpt the Future Foundation for funding for analysis and reporting. We thank Catalina Gomez for advice on an early draft of this report, and for assistance with upcoming comparisons of our 2002 density estimates to her more recent density estimates from the same study site.

REFERENCES

Buckland ST, Anderson DR, Burnham KP, Laake JL, Borchers DL and Thomas L (2001) *Introduction to distance sampling: estimating abundance of biological populations*. Oxford University Press, Oxford, UK

da Silva, R. P., and A. R. Martin. 2000. The status of the boto or Amazon river dolphin *Inia geoffrensis* (de Blainville, 1817): a reiew of available information. Paper presented to the Scientific Committee of the International Whaling Commission, Small Cetaceans Sub-Committee.

Fewster RM, Buckland ST, Burnham KP, Borchers DL, Jupp PE, Laake JL and Thomas L (2009) Estimating the encounter rate variance in distance sampling. *Biometrics* 65: 225-236

Gomez-Salazar, C., M. Coll, and H. Whitehead. 2012. River dolphins as indicators of ecosystem degradation in large tropical rivers. Ecological Indicators **23**:19-26.

Huang, S.-L., Y. Hao, Z. Mei, S. T. Turvey, and D. Wang. 2012. Common pattern of population decline for freshwater cetacean species in deteriorating habitats. Freshwater Biology **57**:1266-1276.

Innes S, Heide-Jørgensen MP, Laake JL, Laidre KL, Cleator HJ, Richard P and Stewart REA (2002) Surveys of belugas and narwhals in the Canadian High Arctic in 1996. *NAMMCO Scientific Publications* 4: 169-190

Laake JL and Borchers DL (2004) Methods for incomplete detection at zero distance. In Buckland ST, Anderson DR, Burnham KP, Laake Jl, Borchers DL and Thomas L (eds) *Advanced distance sampling*. Oxford University Press, Oxford, UK

Marques FFC and Buckland ST (2004) Covariate models for the detection function. In Buckland ST, Anderson DR, Burnham KP, Laake Jl, Borchers DL and Thomas L (eds) *Advanced distance sampling*. Oxford University Press, Oxford, UK

Taylor, B. L., M. Martinez, T. Gerrodette, J. Barlow, and Y. N. Hrovat. 2007. LESSONS FROM MONITORING TRENDS IN ABUNDANCE OF MARINE MAMMALS. Marine Mammal Science **23**:157-175.

Thomas L, Buckland ST, Rexstad E, Laake JL, Strindberg S, Hedley SL, Bishop JRB, Marques TA and Burnham KP (2010) Distance software: design and analysis of distance sampling surveys for estimating population size. *J App. Ecol*. 47: 5-14

Vidal, O., J. Barlow, L. A. Hurtado, J. Torre, P. Cendon, and Z. Ojeda. 1997. Distribution and abundance of the Amazon river dolphin ({IInia geffrensis}) and the tucuxi ({ISotalia fluviatilis}) in the upper Amazon River. Marine Mammal Science **13**:427-445.

Williams R, Leaper R, Zerbini AN and Hammond PS (2007) Methods for investigating measurement error in cetacean line-transect surveys. J Mar Biol Ass UK 87:313-320

Table 1 Summary of search effort by strata and line type (km) using distances from the bank.

Stratum	Strip	Line	Strip/Line	Total
Confluence*	1.3		9.4	10.7
Island	25.1	2.1	105.5	132.7
Lake	17.3	12.9	8.8	39.0
Channel	78.2		0.8	79.1
River	20.4	12.3	97.9	130.6
Tributary	25.3			25.3
Total	167.7	27.3	222.3	417.3

*Within 150m of a junction of two water types

Table 2 Summary of sightings by strata and platform (no truncation).

Stratum	All front sightings				Sightings seen by rear only			
	Inia	Sotalia	Unidentified	Total	Inia	Sotalia	Unidentified	Total
Confluence*	*4*	*2*		*6*				
Island	19	13		32	12	4	1	17
Lake	24	36	2	62	12	3		15
Channel	12	14	2	28	3	2		5
River	6	23		29	5	4		9
Tributary	4	7		11	1	1		2
Total	65	93	4	162	33	14	1	48

*These sightings have been excluded from the analyses and totals

Table 3 Frequency of group sizes (no truncation).

Group size	All front sightings			Sightings seen by rear only			All unique sightings		
	Inia	Sotalia	Unidentified	Inia	Sotalia	Unidentified	Inia	Sotalia	Unidentified
1	49	29	3	23	9	1	72	38	4
2	14	43	1	8	4		22	47	1
3	1	11		2	1		3	12	
4		7							
5	1	3					1		
Total	65	93	4	33	14	1	97	97	5

Table 4 Summary of the fitted models; truncation distance (w), the number of groups detected for each platform (n_i; i=1=front, i=2=rear), number of duplicates (m), the probability of detection on the trackline for the front platform ($p_1(0)$) and the overall probability of detection for the front platform p_1.

Sighting type	w (m)	n_1	n_2	m	DS model	MR model	$p_1(0)$	p_1
Line	200	127	131	92	HZ	Species + Size	0.660 (0.08)	0.338 (0.15)
Strip	100	20	21	17	Strip	Petersen	0.810 (0.11)	0.810 (0.11)

Table 5 Group and animal density estimates by strata

Strata	Line type	L (km)	Sighting type	n1	Dg (groups/km^2)	D (animals/km^2)
Island	Line	116.50	Line	26	1.65 (0.32)	2.91 (0.32)
	Strip	1.97	Strip	0	0	0
	Line/Strip	14.22	Line	1	1.85 (0.34)	3.26 (0.35)
			Strip	4		
	Total	132.69		31	**1.65 (0.25)**	**2.91 (0.26)**
Lake	Line	30.98	Line	47	11.22 (0.23)	19.76 (0.23)
	Strip	2.34	Strip	1	2.64 (1.14)	4.64 (1.14)
	Line/Strip	5.68	Line	3	6.66 (0.25)	11.74 (0.25)
			Strip	2		
	Total	39.00		53	**10.41 (0.17)**	**17.69 (0.17)**
Channel	Line	22.60	Line	13	4.25 (0.37)	7.50 (0.38)
	Strip	39.25	Strip	4	0.63 (0.43)	1.11 (0.43)
	Line/Strip	17.21	Line	5	3.34 (0.31)	5.89 (0.32)
			Strip	2		
	Total	79.06		24	**2.75 (0.34)**	**3.98 (0.34)**
River	Line	123.99	Line	27	1.61 (0.30)	2.84 (0.31)
	Strip	0.00	Strip	0	0	0
	Line/Strip	6.57	Line	0	0.63 (1.04)	1.10 (1.04)
			Strip	1		
	Total	130.56		28	**1.57 (0.26)**	**2.75 (0.26)**
Tributary	Line	1.86	Line	2	7.97 (0.15)	14.05 (0.15)
	Strip	18.10	Strip	6	2.05 (0.27)	3.61 (0.28)
	Line/Strip	5.36	Line	3	5.52 (0.06)	9.72 (0.08)
			Strip	0		
	Total	25.32			**3.72 (0.11)**	**5.67 (0.12)**

Figure 1 Search effort

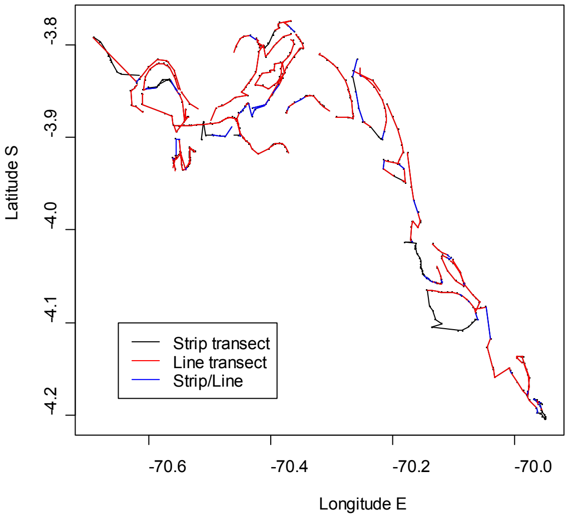
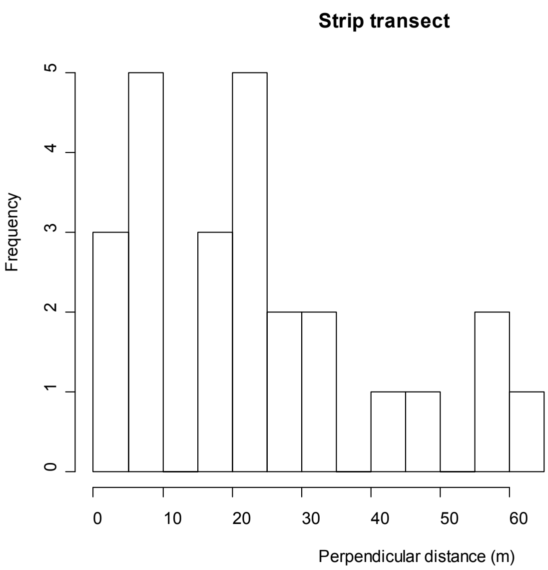


Figure 2 Distribution of perpendicular distances by line type for both platforms combined.



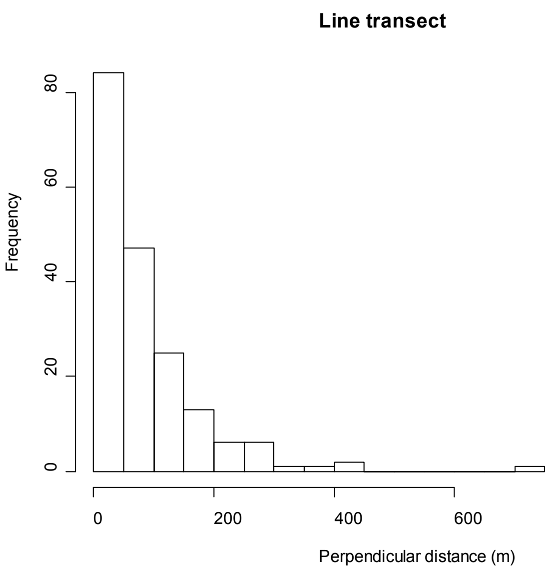
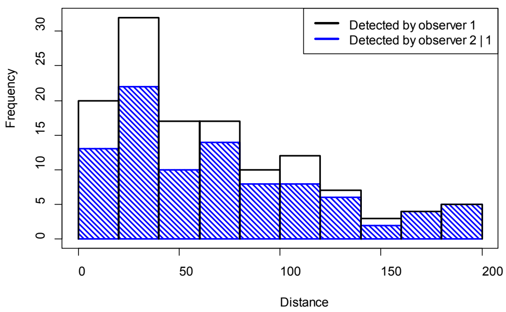


Figure 3 Line transect sightings for the front platform (observer 1) and the rear platform (observer 2). The shading indicates duplicate sightings.



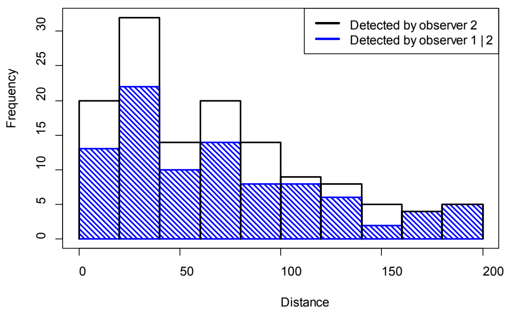


Figure 4 The estimated probability of detection for the front platform given that it was detected by the rear platform; dots indicate individual estimates given the covariates and the line indicates the value averaged overall covariates.

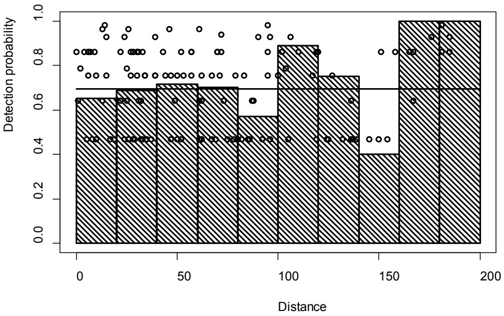


Figure 5 Fitted DS model overlaid onto the scaled perpendicular distance distribution of front line transect sightings.

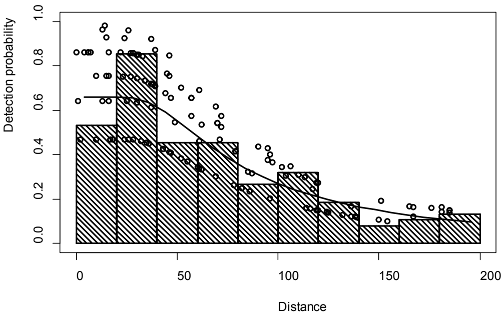


Figure 6 Location of the front sightings as the proportion of the distance from the port shore to river width; sightings close to the port shore will be close to zero and sightings close to the starboard shore will be close to one.