Group IV Humpback Whales: Abundance estimates from aerial and land-based surveys off Shark Bay, Western Australia, 2008

SHARON L. HEDLEY*, JOHN L. BANNISTER† AND REBECCA A. DUNLOP‡

Contact email: sharon.hedley@btinternet.com

ABSTRACT

Single platform aerial line transect and land-based surveys of Southern Hemisphere Group IV humpback whales *Megaptera novaeangliae* were undertaken to provide absolute abundance estimates of animals migrating northward along the western Australian coast. The aerial survey flew a total of 28 flights, of which 26 were completed successfully, from 24th June-19th August 2008. The land-based survey was undertaken from Cape Inscription, Dirk Hartog Island, during the expected peak of the whales' northward migration, from 8th-20th July. During the first week of the land-based survey, some double count effort was undertaken to provide information on the numbers of pods missed from the land station. The assumed period of northward migration was 2nd June-7th September. Estimated abundance of northward-migrating whales during that time is 21,750 (95% CI: (17,550-43,000)). This estimate is based on an estimate of relative abundance of surface-available whales of 11,850 (9,550-23,450), and an estimated $g(0)$ of 0.54 (±0.21).

INTRODUCTION

Following increasing reports of humpback whale (*Megaptera novaeangliae*) sightings in winter off the western Australian coast in the early-mid 1970s, aerial surveys of humpback whales during their northward migration were undertaken from Carnarvon, Western Australia (WA) in an area off Shark Bay where aerial spotter and other data from whaling operations were available for the last year of humpback whaling, 1963. Results of those surveys to 1988 (Bannister *et al*., 1991) demonstrated that significantly more whales were seen in the area in the 1980s than in 1963. Further surveys, in 1991 and 1994, demonstrated an annual increase rate of 10.15 ± 4.6% to 1994 (see Bannister and Hedley 2001). In comparison to the estimated population size of 568 at the end of 1963 (Bannister, 1964), the population size in 1994 was calculated to be some 4000-5000 animals (Bannister, 1995).

The 1994 survey results showed that to detect a significant difference in population in future years, at an annual increase of 10%, an interval of three years would be required between surveys, leading to a proposed further survey in 1997. Given funding constraints, that survey took place in 1999, its aim being to provide an estimate of absolute abundance. This aim was more ambitious than its predecessors, from which only a relative index had been obtained. The survey was planned to cover as much of the northern migration period as possible, with flights every other day over a two month period, mid June – mid August. Given the prevailing generally poor weather conditions, only 18 of the 30 planned flights could be flown, of which only 15 were completed. Nevertheless allowing for animals missed while submerged, 1999 population size was estimated as 8200-13600 (Bannister and Hedley, 2001).

Given the disappointing coverage, a further survey was planned to take place as soon as possible over the same period and area, but to include an additional land-based component. That survey took place in 2005; the results are reported in Paxton *et al*. (in press). Unfortunately, although the 2005 survey had been designed with the aim of improving on earlier surveys (which were only able to apply *ad hoc* corrections to adjust for uncertain trackline detection), last-minute logistical changes to the land-based survey in 2005 reduced its effectiveness. In particular, the location of the land-based survey had to be moved northward to a location where, in the event, whales often exhibited 'milling' behaviour rather than directional swimming more typical of migrating animals, and to where the offshore distribution of whales extended far beyond the visual range of the land-based observers.

Given rather equivocal results from the 2005 survey, improvements to the design of the 2008 survey were planned as follows:

1. The aerial survey component was expanded in area to extend offshore coverage (following some experimental work in 2007 to determine the most appropriate survey area).
2. Aerial survey data were collected using a direct data acquisition system.

* *The School House, Denhead, St Andrews, Fife, KY16 8PA, United Kingdom.*
† *Western Australian Museum, Locked Bag 49, Welshpool DC, Western Australia 6986, Australia.*
‡ *Cetacean Ecology and Acoustics Laboratory School of Veterinary Science University of Queensland St Lucia, Qld 4072, Australia.*

3. The land-based component of the survey was expanded to include some double-platform independent observer counts, and thus allow estimation of a correction factor for whales missed by the land-based observers.
4. The location of the land-based platforms was at Cape Inscription, Dirk Hartog Island, Shark Bay. From previous surveys, it was expected that whales passing this location would be more identifiable as 'northward-migrating' and furthermore, that they would pass closer to the shore at this latitude.

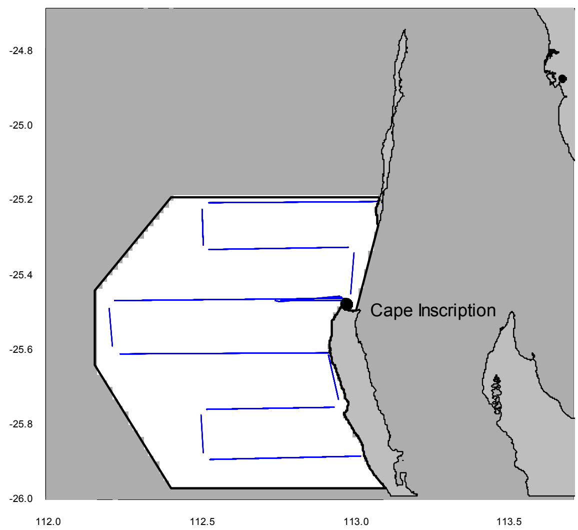
This report details the analysis of data from the 2008 survey, the aerial component of which took place from 24 June-19 August, with the land-based component from 8-20 July. Details of the field methods for the aerial survey are given in Bannister *et al.* (2009, unpublished), and the land-based survey in Dunlop (2008, unpublished).

DATA SUMMARY

Aerial survey

A total of 28 flights were flown; 26 of these were successfully completed and included in the analysis. The survey area and a typical flight path are shown in Figure 1. The approximate length of the two most northerly and two most southerly east-west transects was 45-50km. The extended transects located off the north of Dirk Hartog Island were each approximately 70km in length. In addition, for seven flights which occurred on days on

Figure 1: Survey area for aerial survey, and typical flight path. (Flight 8 on 10th July shown.)



which the land-based survey was also operating, some short transects (of approximately 20km) were flown at the latitude of Cape Inscription. The survey area covered a region of approximately 6570km^2.

Two fixed-wing aircraft were used in the survey: a Partenavia aircraft, fitted with bubble windows (used on 24 flights), and a Cessna 337, with flat windows (used on 4 flights). On each flight, there were two observers, one on each side of the aircraft, and four observers were used in total. Observer participation ranged from flying 24 of the 28 flights (85%) down to 7 (25%).

The first three flights (on 24, 26 and 29 June) were flown in a northerly direction; the remainder were flown in a southerly direction. The latter is preferable since it is in the opposite direction to the whales' migration path but for logistical reasons, this was not possible on the first three flights.

Data collected on each flight included data from a GPS stream (such as time, position and altitude); sightings data (such as angle of declination to the sighting, swimming direction and pod size); and effort data (wind speed, cloud cover, sightability, etc.)

Table 1 details the date, total transect length and number of sightings for each flight. 'NM' sightings are those pods recorded with a northward swimming direction. NM+ sightings additionally include some pods of undetermined direction, randomly allocated to be travelling north in proportion to the sightings of known direction on a given day which were travelling northwards.

FLIGHT	AIRCRAFT	DATE	EFFORT (KM)	NM PODS (AFTER LEFT-TRUNCATION)		NM WHALES (AFTER LEFT- TRUNCATION)		NM+ PODS (AFTER LEFT-TRUNCATION)		NM+ WHALES (AFTER LEFT- TRUNCATION)	
1	Partnv	24/06/08	540	12	(12)	26	(26)	17	(17)	33	(33)
2	Partnv	26/06/08	410	3	(3)	3	(3)	6	(5)	8	(7)
3	Partnv	29/06/08	530	8	(5)	20	(13)	8	(5)	20	(13)
4	Partnv	02/07/08	570	43	(40)	71	(66)	57	(54)	92	(87)
5	Partnv	03/07/08	470	20	(19)	39	(37)	28	(27)	48	(46)
6	Partnv	08/07/08	540	29	(28)	55	(53)	35	(34)	67	(65)
7	Partnv	09/07/08	550	37	(35)	78	(72)	51	(49)	96	(90)
8	Partnv	10/07/08	510	53	(50)	83	(78)	67	(63)	100	(94)
9	Cessna	13/07/08	500	30	(30)	66	(66)	42	(41)	84	(82)
10	Cessna	14/07/08	570	46	(46)	68	(68)	54	(54)	77	(77)
11	Partnv	16/07/08	580	21	(20)	35	(33)	78	(76)	115	(112)
12	Cessna	17/07/08	580	15	(14)	32	(31)	32	(30)	55	(51)
13	Partnv	22/07/08	480	29	(25)	60	(49)	68	(62)	115	(101)
14	Partnv	23/07/08	480	37	(32)	70	(59)	56	(51)	95	(84)
15*	Partnv	24/07/08	190	7	(6)	9	(7)	11	(10)	13	(11)
16	Partnv	29/07/08	460	32	(30)	48	(44)	58	(56)	79	(75)
17	Partnv	02/08/08	490	15	(12)	25	(20)	37	(34)	52	(47)
18	Partnv	06/08/08	440	15	(15)	28	(28)	23	(23)	36	(36)
19	Partnv	08/08/08	460	7	(7)	13	(13)	14	(14)	23	(23)
20	Partnv	09/08/08	470	15	(13)	21	(19)	27	(24)	38	(35)
21	Partnv	10/08/08	470	23	(21)	43	(41)	28	(26)	48	(46)
22	Partnv	12/08/08	480	12	(12)	16	(16)	20	(19)	26	(25)
23	Partnv	13/08/08	480	17	(16)	28	(26)	26	(25)	38	(36)
24	Cessna	14/08/08	440	5	(5)	8	(8)	8	(8)	12	(12)
25	Partnv	15/08/08	400	12	(12)	21	(21)	23	(23)	35	(35)
26	Partnv	16/08/08	470	16	(16)	24	(24)	26	(26)	35	(35)
27*	Partnv	18/08/08	190	4	(4)	7	(7)	8	(8)	14	(14)
28	Partnv	19/08/08	470	8	(8)	11	(11)	12	(12)	15	(15)
	TOTAL		13,220	571	(536)	1008	(939)	920	(876)	1469	(1387)

Table 1 Summary of aerial surveys. Flights marked with an asterisk were aborted and their data excluded from the analysis. Numbers in parentheses are the numbers of pods/whales after left-truncation of perpendicular distances at 260m.

Land-based survey

The land-based survey took place from Cape Inscription, on the northern end of Dirk Hartog Island, from 8-20 July. Survey effort was scheduled for 9 hours each day; 7 full days were completed and three partial days (of 6, 7 and 2.5 hours respectively), with no effort possible on 11 and 20 July. During the first survey week (8-13 July), 5 hours of double-platform (independent observer) data were collected on each day with suitable survey conditions (25 hours in total). During the second week, reduced personnel resulted in it only being feasible to conduct single-platform survey; these observations were augmented by 'focal follows' (i.e. each surfacing of a detected pod recorded until out of visible range) without disruption to the sightings survey.

DATE	EFFORT (HOURS)	DOUBLE PLATFORM EFFORT (HOURS)	NM PODS	NM PODS DIST270 TRUNCATED AT 12KM	NM PODS DIST270 TRUNCATED AT 12KM (PODS WITH NO DIST270 INCLUDED)	NM+ PODS	NM+ PODS DIST270 TRUNCATED AT 12KM	NM+ PODS DIST270 TRUNCATED AT 12KM (PODS WITH NO DIST270 INCLUDED)
08/07/08	9	5	28	23	27	36	25	31
09/07/08	9	5	14	6	14	15	6	15
10/07/08	9	5	19	11	17	25	13	22
11/07/08	0	0	0	0	0	0	0	0
12/07/08	9	5	23	10	18	24	10	19
13/07/08	6	5	32	11	22	43	6	30
14/07/08	6	0	13	6	11	16	8	12
15/07/08	7	0	17	7	13	20	33	15
16/07/08	9	0	42	31	42	46	15	46
17/07/08	9	0	23	13	20	23	0	20
18/07/08	2.5	0	15	0	15	16	11	16
19/07/08	9	0	16	11	16	16	0	16
20/07/08	0	0	0	0	0	0	0	0
TOTAL	84.5	25	242	129	215	280	127	242

Table 2 Summary of land-based survey effort and humpback whale pod sightings. Sightings shown for NM and NM+ pods. 'Dist270' is the perpendicular offshore distance.

As emphasised in Dunlop (2008, unpublished), the conditions for the land-based survey were far from ideal. The terrain is rugged and exposed, with virtually no facilities at the site. Moreover, for the survey itself, the site was low – the highest accessible point being just 25.5m above sea level. In the event, a large proportion of the whales migrated past this point some distance from the shore, resulting not only in a high proportion of whales being missed, but also in difficulties obtaining theodolite fixes required for tracking of pods and accurate distance estimation. Dunlop (2008, unpublished) recommended that 12km be used as the maximum truncation distance for pods sighted from land, although beyond about 8km, whales were sighted on the horizon so even beyond his distance, recorded distances may be unreliable. The implications of these inaccuracies for this analysis have not been fully considered here, but some potential issues are noted in the Discussion.

The matching process (undertaken by R. Dunlop) was certainly severely hampered by the distance inaccuracies, but is assumed to have been completed without error in this report (i.e. no account is taken of incorrect duplicate identification). Data collected by the land-based teams included Pod ID, bearing, distance and angle to the pod at time of detection, swimming direction, pod size, and perpendicular distance offshore (rarely observed but calculable from a second fix to the pod having passed the 'abeam' line from the land-based platform). For the double-platform data, weather and sightability conditions were also recorded, along with an assessment of duplicate status. A summary of the land-based survey data is shown in Table 2. The number of NM and NM+ pods sighted is given, together with two further datasets: (1) the number of sightings after truncation at 12km offshore (and excluding pods for which no offshore distance was available; and (2) the number of sightings after truncation at 12km offshore (and including those pods with no offshore distance).

In addition to the survey data, a total of 22 focal follows were conducted during the land-based survey. Details of these data are given in Dunlop (2008, unpublished), where they were also analysed to give an average speed of northward-migrating travel of 5.56kmh^{-1} (±0.31). This figure was used in the analyses presented here to estimate the rate of passage of pods through the survey area (and hence convert 'snapshot' aerial survey estimates to daily numbers of pods).

ANALYSIS METHODS

Overview

The survey objective was to estimate the absolute abundance of northward-migrating humpback whales off Shark Bay. The aim of the aerial survey component was to estimate the number of whale pods seen on a given flight. This number would then require a correction so that it corresponded to the number of pods passing through the area during a given time, say, per day. Such a correction factor would depend on the whales' speed of travel during their northward migration. Without further adjustment, the number of pods per day would be an underestimate of the true number, since it is known that aerial line transect surveys for humpback whales are negatively biased. Broadly speaking, uncorrected estimates only estimate the number of whales *at the surface* and thus available to be seen; in addition to this 'availability' bias, not all whales at the surface are detected, leading to so-called 'perception' bias (Marsh and Sinclair, 1989).

The aim of the land-based survey component was threefold: (1) to provide an estimate of absolute abundance of northward-migrating humpback whale pods during the two weeks of the aerial survey (and thus allow calibration of the corresponding aerial estimates); (2) using the focal follow data, to provide estimates of whale migration speed; and (3) to provide estimates of mean pod sizes (since it was expected that these would be underestimated from the aerial survey).

Combining the results from the two components, estimates of the absolute number of northward-migrating whales passing through the survey area for each day of the aerial survey may be obtained. Fitting a model to these estimates (to allow prediction of the number of whales passing through the area on non-survey days, including those at the very beginning and end of the expected period of northward migration), and integrating the fit throughout the migration period, yields an estimate of absolute abundance of northward migrating whales.

Modelling the aerial survey data to obtain relative density estimates

Note that in what immediately follows, 'density' refers to 'relative density', since no account for perception nor availability bias has been made (i.e. in this section, $g(0)$ is assumed to be equal to one).

For each flight, pod density is estimated using a spatial generalized additive model (GAM) similar to the 'count model' of Hedley and Buckland (2004). The response variable of the model is the number of pod sightings per 'segment' of the transect, where the segment length must be specified but should be selected such that sighting conditions (and geographic location) do not change appreciably within a segment. An offset variable is included in the model to account for differences in estimated probabilities of detection within each segment, and consequential potentially different effective search areas of the segments. The offset is estimated using multiple covariate distance sampling – conventional single platform line transect estimation but with the ability to include covariates (such as sea state) in the scale parameter of the detection function (Marques and Buckland, 2003).

With a logarithmic link function, the general form of a GAM of this type may be written

$$E[n_i] = \exp\left\{ \log\left(2\, l_i w \cdot \hat{p}_i\right) + \sum_k f_k(z_{ik}) \right\},$$

where $E[n_i]$ is the expected number of sighted pods in the i^{th} segment and $\mathrm{Var}[n_i]$ is assumed to be proportional to this; l_i is the length of segment i; w is the perpendicular (right-) truncation distance; $\hat{p}_i$ is the estimated probability of detection of a pod in segment i; z_{ij}, $j=1,\ldots k$ denotes the value of the j^{th} (spatial) covariate in the i^{th} segment; and the f_k are (smooth) functions. Extending this form, it is feasible for a function f_j to depend on more than one covariate (e.g. $f(lat_i, lon_i)$), and/or for the covariate to be temporal (e.g. *day*).

Hedley and Buckland (2004) suggested that variance from a spatial model of this type may be estimated using an appropriate resampling scheme such as a non-parametric or parametric bootstrap. In practice, these bootstrapping techniques frequently give biased results when smoothing models. Wood (2006, p246-7) proposed an alternative approach which can be much simpler to implement, and appears not to suffer from the bias often associated with the bootstrapping approaches. This approach uses a 'prediction matrix' to map the model parameters to the predictions of the linear predictor, in conjunction with simulation from the posterior distribution of the parameters. The analysis in this report uses Wood's (2006) approach, conditioning on the estimated smoothing parameters.

The model form given above includes as a predictor variable, an estimate, $\hat{p}_i$, of the probability of detection. Whilst a bootstrapping approach could be implemented to include variance in this estimate, as noted above bootstrapping spatial models often gives unstable and biased results. An alternative method of propagating the uncertainty in this estimate has been implemented in this analysis. The idea is currently being developed

(Bravington *et al*., in prep.). A matrix of first derivatives of log(*eff.area$_i$*)=$\log\left(2 l_i w \cdot \hat{p}_i\right)$ with respect to the parameters of the detection function is included in the linear predictor. The Hessian matrix from the likelihood maximization of the detection function describes the local curvature of the fit associated with the parameter estimates; its value(s) are used in the model fitting process also (as a prior on the variance of a parameter to be estimated by the spatial model). Algebraic details are given in the Appendix.

Estimating mean pod size

Results from other studies have shown that aerial survey pod size estimates can be negatively biased, since the animals are in view only for a relatively short period of time. In contrast, some pods sighted from the land station were tracked for over an hour. Prior to this analysis therefore, it was expected that the 'best' estimate of mean pod size would be derived from the land-based data. Dunlop (2008, unpublished) noted problems with tracking pods on this survey because of the high numbers of pods seen far offshore, but concluded that the best estimate of mean pod size from the land-based survey was 1.717 (±0.088).

Estimating abundance from the land-based survey data

Within the visible range of the land-based observers, say up to 12km offshore, the number of northward-migrating whales passing the land station per watch period (where a 'watch' is defined as a 3 hour period within a day, say) gives an estimate of their rate of passage. Using the double-platform data from the first survey week, we use a logistic regression approach based on that used by Buckland *et al*. (1993) for grey whales (*Eschrichtius robustus*) to estimate the proportion of whales missed, and hence to enable correction factors (and their standard errors) for the number of pods seen to be estimated.

Three correction factors are estimated, depending on the mode of survey operation at the time (i.e. [§]Platform 1 only, [**]Platform 2 only or Double Platform). It is assumed that the probability of detection of a pod from one platform is independent of whether it is detected from the other, and independent of whether other pods are detected by either platform.

The counts from each watch are then adjusted according to the mode of survey operation. Summing, and standardizing for different hours of effort, daily estimates of pod abundance are calculated. The estimates correspond to the survey region in view from the land-based station only.

RESULTS

Use of the aerial data

Prior to analysis, transect line lengths were calculated from the GPS positional data using R code adapted from functions written in Visual Basic (by J.L. Laake, NMML). Corresponding formulae are given in Zwillinger (2002). Heading angles were corrected for aircraft drift angle, and perpendicular distances (x) to sightings were calculated using the following simple tangent formula (e.g. Pike *et al*., 2008):

$$x = h(\tan(90-\theta))\sin(\phi),$$

where h is altitude; θ is declination angle to the sighting,; and φ is drift-correcting heading angle.

During the aerial survey, the swimming direction of sighted pods was recorded where possible. Since the objective of the survey is to obtain estimates for the northward-migrating component of the population only, then the swimming direction is critical. Out of 855 pods with either a swimming direction recorded, or designated as 'milling', then 571 (67%) of these were recorded as travelling northwards (where NE and NW were classified as North). In total, 1357 humpback (including 'possible' humpback) pods were recorded whilst on effort and 42% of these were recorded as travelling northwards. As in Paxton *et al*. (in press), humpbacks with no direction recorded (and not milling), were randomly allocated a swimming direction according to the relative proportions of directions observed on a given flight. This increased the sample size considerably to 920 northward-migrating whales (seen on effort). Hereafter, we analyse the data for whales recorded as travelling north (NM whales) separately from a dataset of NM whales augmented by sightings with unknown swimming direction, but randomly allocated to be travelling northwards (NM+ whales).

[§] Termed 'Car' platform in Dunlop (2008, unpublished).
[**] Termed 'Bush' platform in Dunlop (2008, unpublished).

Detection function estimation: aerial data

Two aircraft were used on the aerial survey: the Partenavia, fitted with bubble windows, and the Cessna, with flat windows. Angles of declination taken from each aircraft suggested that strips of about 80m (40m either side of the trackline) and of about 260m were obscured from the view of observers immediately beneath the Partenavia and the Cessna respectively. Histograms of perpendicular distances suggested that some pods were being missed outwith this strip for the Partenavia, perhaps because it was uncomfortable for the observers to look down at such an angle. This problem was alleviated by extending the left-truncation distance to 260m for both aircraft; thus about 6% of the sightings were excluded from further analysis (see Table 1).

Initial exploratory analyses of the NM aerial line transect data were conducted in *Distance* v5.0 (Thomas *et al.*, 2006), and model selection for both NM and NM+ whales was based on these analyses. Potential factors or covariates included *Cloud cover, Sightability, Side of Aircraft* (Port/Starboard), *Sea state, Wind speed, Observer, Pod size* and *Aircraft*. The detection function was modelled as a function of perpendicular distance, and these variables were considered for inclusion via the scale parameter of this function (either a hazard-rate or a half-normal form). The perpendicular distance data were right-truncated at 3.0km for NM whales and 4.5km for NM+ whales. A stepwise forward selection procedure (starting with a model containing perpendicular distance only) based on Bayes' Information Criterion (BIC) was used for model selection.

For both NM and NM+ pods, the model selected by BIC alone would have included *Pod size*. However the fitted detection function from such a model was such that estimated probability of detection decreased as pod size increased, counter to expectation. For NM+ pods, the BIC also suggested a model including *Sightability* was better than a perpendicular-distance-only model. Similarly to pod size, however, probability of detection was estimated to be lower in 'Excellent' conditions than in 'Good' and 'Poor' conditions. The other covariates were not found to significantly improve upon a perpendicular-distance-only fit, and so in the absence of an explanation for the relationship between detectability and pod size, or between detectability and sightability, half-normal models of perpendicular distance only were fitted to both the NM and the NM+ data. Fitted detection functions are shown in Figure 2. Estimated effective strip half-widths were 2.05km (± 0.088) and 2.46km (± 0.084) respectively.

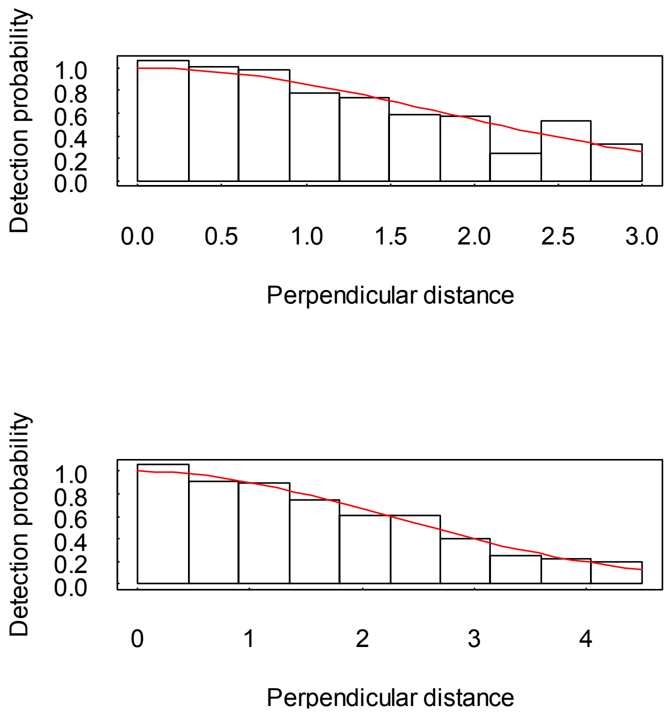


Figure 2 Fitted detection functions (half-normal models) for aerial survey data. Perpendicular distances in km.

NM pods: left panel. NM+pods: right panel.

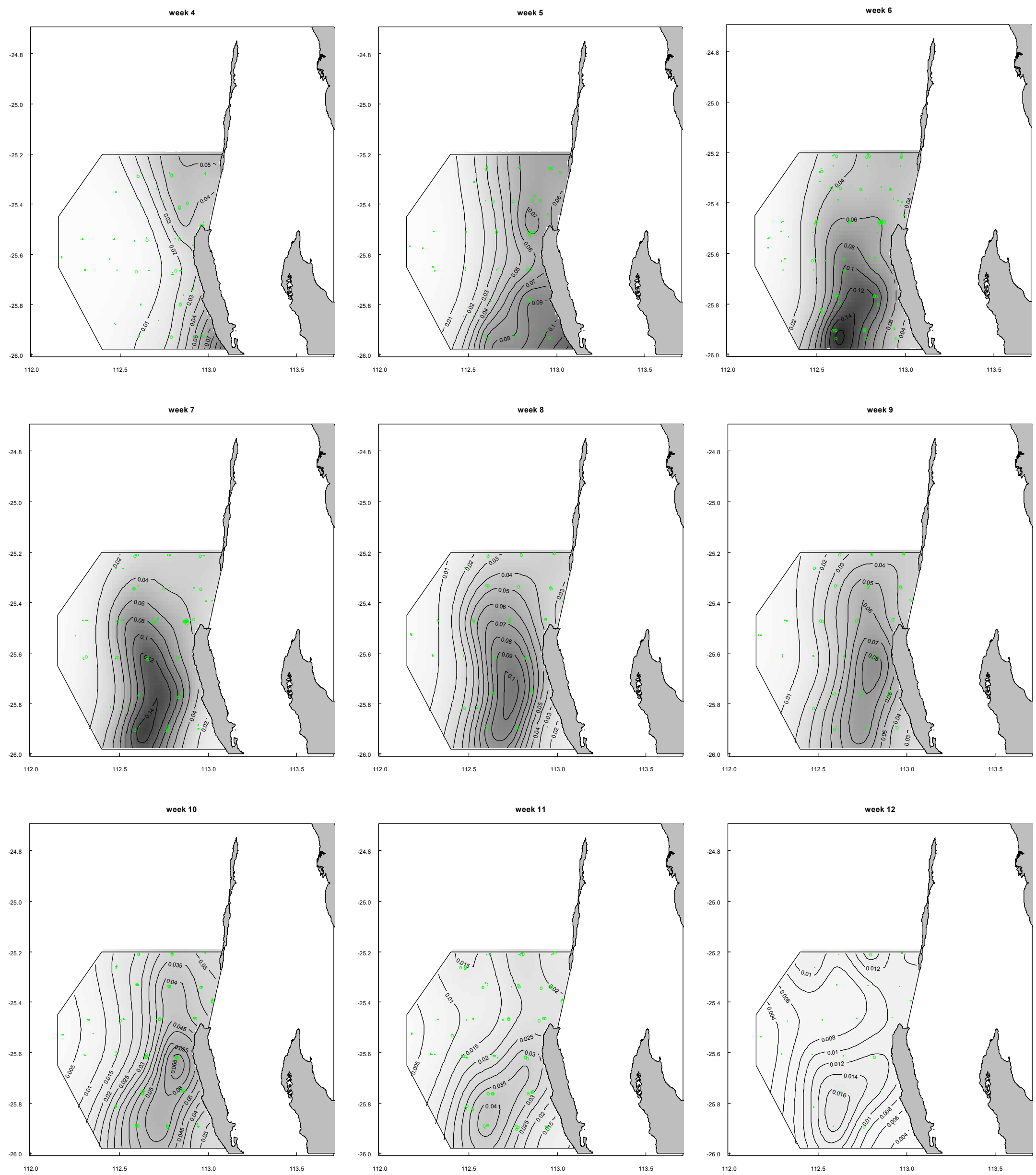


Figure 3 Estimated spatial variation in NM pod density throughout the northward migration season, estimated from the aerial survey data. Weeks 1-3 and 13-14, all of which had relatively low densities, not presented here. Green circles/dots represent ‘data’, i.e. weeks with at least one flight conducted (circles represent a segment with at least one sighting; dots represent no sightings in a segment). Week 2 corresponds to the w/c June 9th 2008. Week 12 corresponds to the w/c August 18th 2008.

Mean pod size estimation

As a quick check on the suitability of using the land-based survey estimate for the aerial survey data, a simple spatio-temporal model was fitted to the recorded pod sizes of NM whales from the aerial survey. To reduce the effects of size-bias, only pods detected within 0.7km of the trackline were included in the model. No variation in mean pod size was detected, either spatially or temporally during the period of the aerial surveys.

There was no evidence of ‘size bias’ effects from the aerial survey data, although as noted above, any effect of pod size on detectability appeared to be in the ‘wrong’ direction. Mean pod sizes from the aerial data were estimated as 1.80 (±0.043) and 1.64 (±0.032) for NM and NM+ whales respectively (cf. from the land-based survey 1.717 (±0.088)). The lower estimate for NM+ whales is not surprising, since this data set includes pods for which a swimming direction was not recorded, so presumably pod size would be more difficult to ascertain for such pods also (and would tend to be under-estimated). Therefore, in this analysis – for both NM and NM+ estimates – a mean pod size of 1.80 was used.

Spatio-temporal model of the aerial data
Transects covered on effort were divided into segments of length approximately equal to 10 nmiles (18.5km), and the number of pods sighted in each segment was calculated. For each segment, an offset variable was computed as the logarithm of the effective area of the segment, where the effective area is given by twice the segment length multiplied by the estimated effective strip half-width from the detection function estimation described above. Potential spatial covariates were *Latitude*; *Longitude* and *Bottom depth* – sourced from a 1′ by 1′ grid from the U.S. National Geophysical Data Center, NOAA Satellite and Information Service (www.ngdc.noaa.gov/mgg/bathymetry). In addition, *Day* or alternatively, *Week* (where Day 1 – and the first day of Week 1 – was defined to correspond to 2 June, the assumed start of the whales' northward migration period) were potential temporal covariates.

Model fitting and model selection were conducted in the *mgcv* package (Wood, 2008) available in R (R Development Core Team, 2008), with inflated model degrees of freedom to reduce the tendency of generalized cross validation to overfit (Kim and Gu, 2004). A number of forms for the smoothing components of the spatial models were considered, but none of these showed evidence for including *Bottom depth* in the model. GCV score was used to compare models; the final selected model incorporating a tensor product smooth (Wood, 2006) of a two-dimensional thin-plate spline of *Latitude* and *Longitude*, and a thin-plate spline of '*Day'*.

$$\log [E(nsight_i)] = te(Latitude_i, Longitude_i, Day_i) + \log(estimated\ effective\ area_i) + X_i$$

where Var($nsight_i$) was assumed to be proportional to E($nsight_i$), and te is a tensor product of thin-plate spline smooths of *Latitude* and *Longitude, and Day*. The offset variable for the i^{th} observation, log(*estimated effective area*$_i$), was estimated using the effective strip widths estimated from the distance sampling analysis. X is a vector of first derivatives and was used to propagate variance, penalized according to the Hessian of the respective detection function fit. Estimation of tail densities (before the first flight of the season and after the last) was improved by adding two zero counts to the data, one on 2nd June and one on 7th September.

A similar model, with *Week* instead of *Day*, yielded the plots shown in Figure 3, demonstrating how the distribution of whale pods varies during the course of the migration period. At the latitude of Cape Inscription, the estimated pod density as a function of distance offshore (averaged over each week of the aerial survey – i.e. weeks 7 and 8) is shown in Figure 4. These plots indicate that density in week 7 increased gradually with distance offshore to a peak at around 30-35km offshore. During week 8, peak density was over a larger distance, at around 20-35km offshore. In both weeks, estimated density was very low beyond about 60km offshore. Within the region of the land-based station (lower panels of Figure 4), the increase in density with distance offshore was slightly greater (and slightly more pronounced) during the second week.

Returning to the spatial model fitted above, integrating under the predicted density surfaces for each day within the assumed migration period gave snapshot estimates of abundance for each survey. To convert these estimates into daily estimates, the rate of passage through the survey area was estimated using an average speed of travel of travel of of 5.56kmh^{-1}. The latitudinal width of the survey area was 86.7km, hence the snapshot estimates were multiplied by a correction factor equal to (5.56 x 24)/86.7 to convert them to daily estimates. (Note that in the present analysis, the estimated variance in speed of travel was not incorporated in the variance of the final abundance estimates.) Multiplying by the estimated mean pod size resulted in estimates of whale abundance, uncorrected for availability and detection bias (Figure 5).

Land-based survey
Noting that sightings from the aerial survey extended far beyond the visible range of the land station, it was clear that an 'abundance' estimate from the land-based survey, even for the two weeks of its duration, would only represent a proportion of the migrating population. In this section, the abundance estimate calculated corresponds to migrating animals passing within 12km of the shore. To use this estimate for calibration of the aerial estimates above requires abundance to be estimated for a corresponding region from the aerial survey (see 'Calibration of aerial survey' below).

To estimate the number of pods missed within 12km offshore during the land-survey, the double count data collected during the first week of that survey were used. The method of Buckland *et al*. (1993) was slightly modified in that the generalized linear model they proposed was replaced by an equivalent GAM formulation. The additional flexibility of the latter seemed to assist in convergence of the model (the offset was estimated iteratively) although this was not fully investigated. In order to obtain a reasonable sample size, this model was

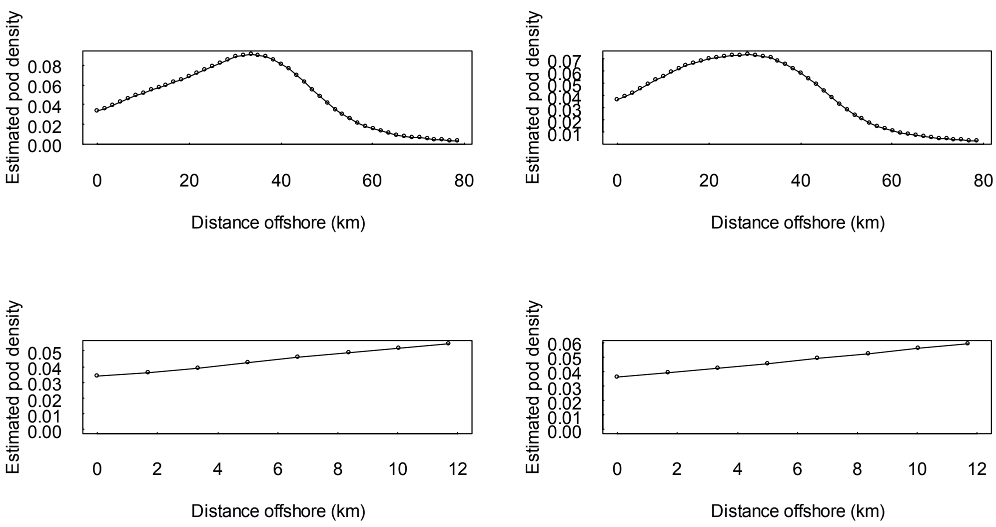


Figure 4 Estimated pod density as a function of distance offshore (from Cape Inscription). Left panels for week 7 (w/c 8th July 2008); right panels for week 8 (w/c 15th July 2008). Upper panels show the estimated density from the shore to the western edge of the survey area; lower panels give this for the first 12km offshore only.

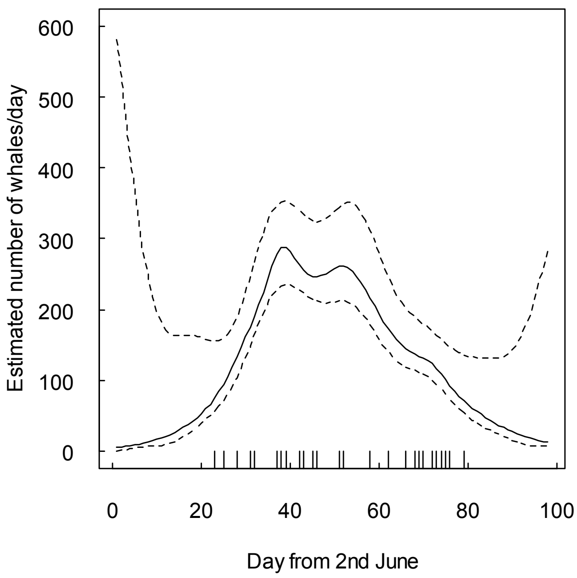


Figure 5 Estimated whale abundance throughout the migration period from spatial modelling of aerial survey data. Dashed lines shows 95% percentile intervals obtained by simulating from the posterior distribution of the parameters of the fitted model. The intervals shown include variance in mean school size, but not in whales' migration speed. Rug plot (long ticks) along the x-axis shows days during this period on which flights were completed.

fitted to NM+ data. After truncation at 12km, the number of pods seen by at least one land platform was 48; distance offshore and encounter rate were included in the final model. With no truncation, the number of pods seen increased to 73, the same covariates plus wind speed were included in the resulting model. Estimated correction factors for the three modes of operation of the land survey (single platform – Car, single platform – Bush, and double platform) are shown in Table 3.

The number of pods seen on each watch period of the land survey was then adjusted according to the correction factors in Table 3. Since there was some daily variation in the number of hours of survey effort, the estimates were also standardized by effort. Using the land-based mean pod size estimate of 1.717, estimates for NM whales corrected and standardized to 9 hours per day are shown in Figure 6. Data from 18th July, on which day there were 2.5 hours of effort, were excluded from the analysis. The total estimated number of pods was 180 (totalling 306 whales).

Truncated at 12km			Untruncated		
Missed by both	Missed by Car	Missed by Bush	Missed by both	Missed by Car	Missed by Bush
1.053 (±0.042)	1.173 (±0.046)	1.324 (±0.053)	1.079 (±0.034)	1.267 (±0.040)	1.425 (±0.045)

Table 3 Estimated correction factors for numbers of pods missed from the land station.

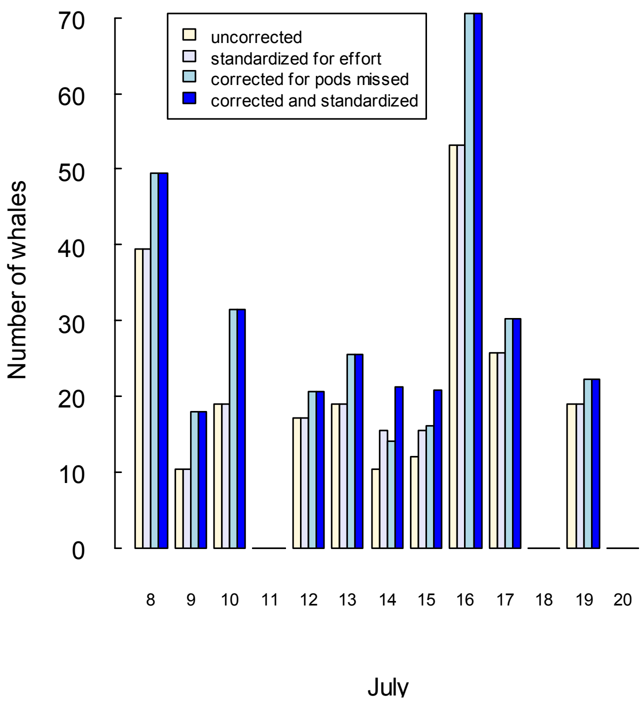


Figure 6 Counts of number of whales passing the land station within 12km of the shore. ‘Uncorrected’ estimates are the raw counts; ‘standardized for effort’ adjusts the estimates to correspond to 9 hours of effort; ‘corrected for pods missed’ uses the correction factors in Table 3 (truncated at 12km) to adjust the counts.

Calibration of aerial survey

Using estimates of pod density at the latitude of Cape Inscription from the spatial model fitted to the aerial survey data, the daily proportion of pods in the survey area that were within 12km of the shore was estimated. For the days of the land survey, the mean proportion was 0.152 (±0.021). Having standardized the number of pods counted from the land to counts corresponding to 24 hours, the number of pods passing the land station each day within the survey area was estimated by adjusting the standardized counts according to the daily proportion of pods that were within 12km. Comparing these estimates to the daily estimates from the aerial survey, and taking the mean, gives estimates of '$g(0)$' for the aerial survey, which account for both perception and availability bias. For NM whales, this was estimated as 0.54 (±0.21). For NM+ whales, it was 0.79 (±0.30). These estimates seem rather high compared to prior expectation based on previous analyses. This could be because, for a large number of whales seen from the land station, no offshore distance was recorded (and thus these pods were excluded from the analysis). Therefore similar estimates were computed, truncating at 12km those sightings with an offshore distance but also included all sightings with no distance recorded. This would most certainly mean that 'too many' pods were included in the land counts, especially since one of the main reasons for a missing offshore distance was difficulty in acquiring two theodolite fixes of the same pod. Nonetheless, it is useful as a sensitivity for this analysis, since Dunlop (2008, unpublished) noted that even beyond 8km, whales were sighted on the horizon. Therefore at least some of the pods with missing distances would be expected to be within 12km offshore. Estimates using these data (and the untruncated correction factors from Table 3) were 0.30 (±0.10) and 0.43 (±0.17). The resulting abundance estimates are shown in Table 4.

Land data, truncated at 12km	NM whales			NM+ whales		
	relative	$g(0)$	Absolute	relative	$g(0)$	absolute
Missing distances excluded	11,850 (9,550-23,450)	0.54	21,750 (17,550-43,000)	14,450 (11,900-23,700)	0.79	18,350 (15,100-30,050)
Missing distances included		0.30	40,300 (32,500-79,700)		0.43	33,850 (27,900-55,500)

Table 4 Estimates of abundance for NM and NM+ whales. The large difference between rows depends on what portion of the land data are used in the calibration of the aerial survey estimates. 'Relative' estimates are uncorrected estimates from the aerial survey; 'absolute' estimates are those corrected by '$g(0)$' estimates from the land-aerial calibration. Numbers in parentheses are 95% percentile intervals; these do not include variance in $g(0)$.

DISCUSSION

The estimates presented in Table 4 are very different, significantly so for the two rows of data which represent different subsets of the land-based data. The land survey was not particularly successful in providing a suitable 'calibration' for the aerial survey estimates, i.e. one that accounted for bias due to a lack of availability of diving pods and due to pods at the surface being missed. This is primarily due to the high proportion of animals that were beyond the range of the land-based observers, and so the overlap between the aerial survey – already for only a few days – was also spatially limited. Additionally, there may be some issues related to the different relative abilities of the aerial and land-based survey to identify the direction of a sighted pod. The land-based survey, for pods sighted sufficiently closely for tracking purposes, recording direction would have been straightforward whereas for the aerial survey, determination of swimming direction would generally have been based on fewer cues over much shorter periods of time in view.

The primary objective of the 2008 survey was to obtain an estimate of absolute abundance of northward-migrating whales. Whilst we can be reasonably confident about the relative estimates presented in Table 4, there is wide variation in the absolute estimates as a result of very different estimates of g(0). *A priori*, from previous analyses and studies elsewhere, estimates in the range 0.3-0.4 or so, would probably have been expected, with this $g(0)$ correcting for both availability bias and perception bias. It is therefore necessary to investigate further the reasons for the apparently high $g(0)$s reported here. The estimation method used by Paxton *et al*. (2005) estimated an 'availability curve' indicating the true (relative) density of pods with distance from shore. Within the region of the land-based observers, this showed a steady increase in density with distance offshore, up to a peak at around 10km. The detection function fitted to the distances offshore (using the land-based data) showed a very steady decrease in detectability with distance, based on a half-normal detection function. Differences between the two curves were used to correct the counts from the land-survey for pods missed from the land, and then *g*(0) was estimated by comparing the aerial abundance in the region with the land-based abundance, over the two-week period of the land-survey in 2005. The correction factor applied to the land data for each day was about 1.5 (C.G.M. Paxton, pers. comm.) The data for the 2008 survey were markedly different from those obtained in 2005. Furthermore, they were very different even between the two weeks of the land survey duration (Figure 7). The improvement to the design of the 2008 survey meant that the estimated number of pods missed from the land was able to be estimated from the double-platform effort during the first week it operated, yielding correction factors by platform operation (see Table 3). The number of pods on which these calculations were based was 73 if the data were not truncated; it decreased to only 48 if the data were truncated at 12km. The estimates of Table 3 appear plausible compared with other studies of migrating populations (e.g. the east coast humpback whale survey of 2004 produced an overall correction of 1.099 (±0.021), but of course, those pods pass much closer to the shore. If the estimates of Table 3 are in fact negatively biased, then the estimates of *g*(0) would be lower (and abundance consequentially higher). Aside from the problems of the offshore distribution of the whales in 2008, then the double-platform land-based approach to estimate the number of pods offshore would be preferable to the aerial-land calibration, since the data would be expected to be more reliable.

An alternative, rather different approach (Barlow *et al.*, 1988) to estimate a *g*(0) correction for availability bias was implemented in Bannister and Hedley (1999) in their analyis of the 1999 survey data:

$$\text{P(being visible)} = (s+t)\ /(s+d)$$

where *s* is the average time a whale stays at the surface; *d* is the average time spent below the surface (i.e. 'deep-diving'), and *t* is the window of time during which an animal is within the visual range of an observer. A range of estimates for the values of *s* and *d* were made based mainly on observational data from experienced humpback whale scientists familiar with Australian animals. A histogram of forward and aft distances was used to gain an idea of the time window, *t.* Ignoring the fact that two aircraft with rather different fields of view were employed on the 2008 survey, a similar histogram of distances to sighted pods is given in Figure 8. This suggests that a maximum sighting 'window' can be estimated as about 8.5km, comprising animals seen ahead (generally up to 5.0km), abeam, and aft (up to 3.5km). However, this window does not appear to be rectangular; so as in Bannister and Hedley (1999), we also compute a correction for smaller sighting windows of 2.5km and 4.5km (estimated from hazard-rate and half-normal fits, respectively, to these data). The focal follow data collected during the 2008 land-based survey were used to provide estimates of *s* and *d* (Dunlop, 2008, unpublished) of 405s and 246s. Average speed during the aerial survey was 132knots. Estimates of *t* using the three different-sized windows are thus 125s, 37s and 66s, giving correction factors of about 0.81, 0.68 and 0.72 respectively – again, much higher than from previous analyses. The *g*(0) estimates from the dive time approach (Barlow *et al.,*1988) are fairly insensitive to quite large changes in window-width (although the extreme of an 8.5km window giving a *g*(0) of 0.81 can probably be discounted). The other *g*(0) estimates obtained using this approach (0.68 and 0.72) are comparable to that from the analysis of NM+ whales in this analysis (0.79). The former do not account for perception bias, however, so would be expected to be higher than that from the combined survey approach. The estimate of 0.54 (±0.21) using NM whales only is thus perhaps more credible.

A second objective of the 2008 survey was to compare results with the 1999 and 2005 surveys. Previous analyses had estimated relative abundance of whales over a similar migration period to that assumed here as 3,441 for 1999 (Bannister and Hedley, 2001) and about 6,030 for 2005 (Paxton *et al*., in press) – an estimated increase rate of 9.8% per annum. The estimate of 11,850 presented here would represent an implausible rate of increase of 14.7% from the 1999 estimate, although this becomes marginally less implausible (12.5%) based on the 2005 estimate alone. Paxton *et al.* (in press) retrospectively applied a correction from their paper to the 1999 estimate to estimate absolute abundance of northward-migrating humpback whales as 11,500 (95% CI 9,200-14,300) which fell within the range of 8,207-13,640 broadly estimated by Bannister and Hedley. This compared with 22,500 (10,000 – 72,200) from the 2005 survey. (Note: The estimate of 22,500 was not considered the 'best' estimate of abundance by Paxton *et al*. (in press) since they considered that extrapolation beyond the last flight of the aerial survey was unreliable due to a presumed 'second pulse' in the migration curve. However, the migration curve fitted in Paxton *et al.* (2005) could be re-fitted with an assumed simpler form, to represent the 'bell-shape' more typical of migrating populations, and if needed with structural zeros to aid the fit in the tails of the distribution. This would make extrapolation beyond the last survey date more feasible, when – as witnessed by data from the last flight – whales were still migrating northwards. It is not clear to what extent this would reduce the estimate, given that the 2005 fit shows a narrower migration peak than fitted to the 2008 data and a sharp increase in numbers from about 3rd August 2005.) The corresponding estimate from the present analysis is 21,750 (17,550-43,000). It is clear from these results that the differences between estimates hinge strongly on the estimate of a *g*(0) correction factor, and hence at this stage at least, comparisons across years should still probably only be made using the relative estimates.

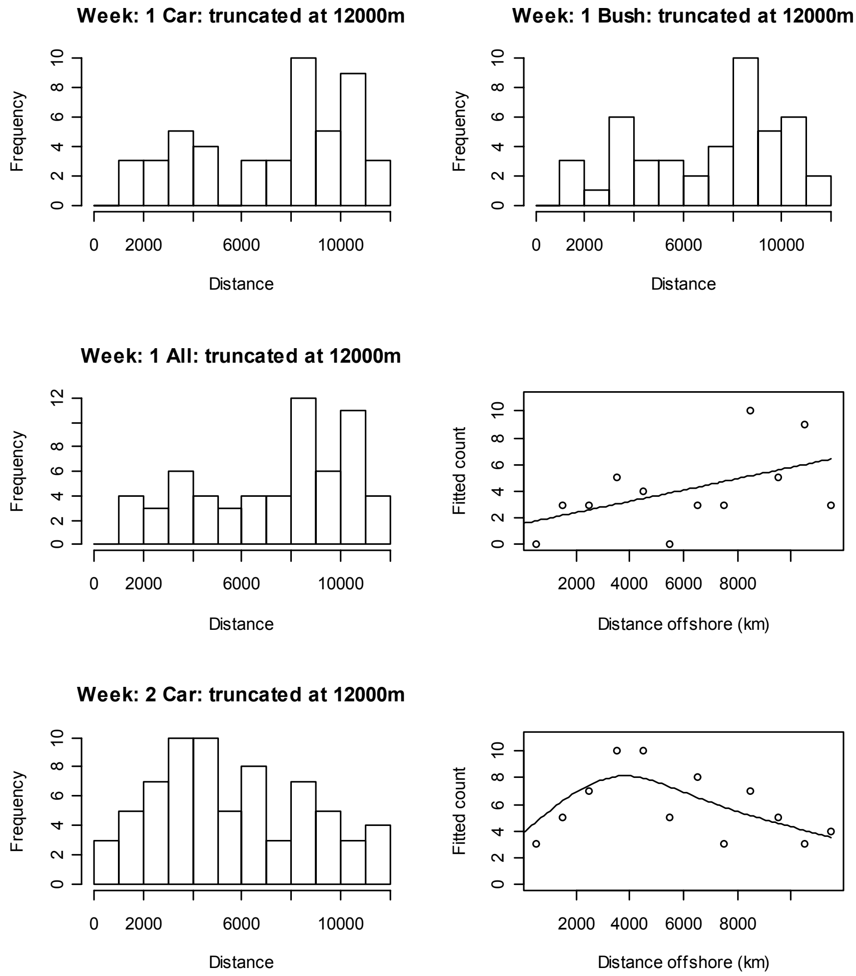


Figure 7 Distribution of NM humpback whale pods with distance offshore, by platform and by week. Data have been truncated at 12km. (During the second week, only the 'Car' platform operated.) Fitted curves are penalized regression splines with smoothing parameters selected by generalized cross validation (Wood, 2006; p130-133).

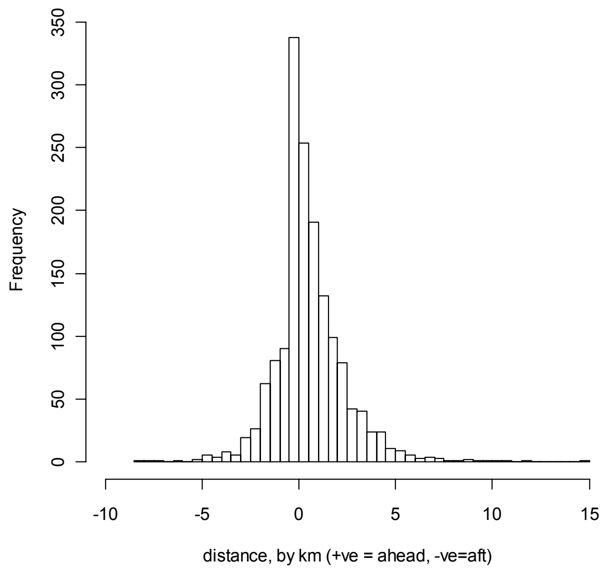


Figure 8 Fore, abeam and aft distances from the aerial survey data.

As a sensitivity test to the spatial modelling approach adopted for analysing the aerial survey data, we also compared the spatial modelling estimates (uncorrected for rate of passage and for g(0)) to those from a conventional line transect analysis in *Distance* (Thomas *et al*., 2006). Data used in the spatial modelling included all on-effort data; only data from the main E-W transects were used in the design-based line transect analysis (as was done previously (Bannister and Hedley, 2005; Paxton *et al.* (in press; results sets 5 and 6). The results are shown in Figure 9. It can be seen that the estimates from the spatial model are quite comparable to those from a standard line transect analysis, the main difference being that variation in encounter rate has been ‘smoothed’ out, as would be expected. This lends weight to the estimates from our analysis, at least those not corrected for *g*(0). Some further investigation remains to be undertaken to untangle the reasons for the apparently high values of *g*(0) reported here, however.

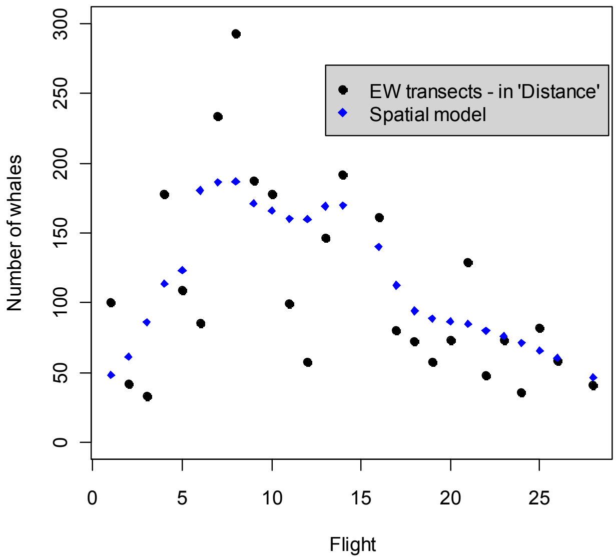


Figure 9 Point estimates of abundance of humpback whales from each flight. Estimates from EW transects are from a conventional line transect analysis in *Distance*; the blue diamonds are from the spatial model fitted in this report.

In conclusion, we propose that the best estimate for NM whales from the 2008 survey is 21,750 (17,550-43,000). The caveat to this is the *g*(0) estimate associated with it of 0.54 (±0.21), which is high compared with previous analyses. Further work is needed to either establish the validity of the latter. Unfortunately, due to the differences in analytical assumptions between the three sets of surveys in 1999, 2005 and 2008, comparisons between these sets are difficult. In 1999, it is not clear (but it is assumed here) that the estimates refer to NM whales. The ‘absolute’ estimates presented in Bannister and Hedley (2001) from that survey were *post hoc* corrections to the relative estimates, and were not directly data-based as are the estimates presented here. In 2005, the effectiveness of the survey was hampered by a late logistical change to the location of the land-based survey, which itself, only operated in single-platform mode. The resulting estimates depend on classification of swimming direction (as here) but it is thought that many of the whales in the area were not migrating, with a high proportion of pods ‘milling’. Furthermore, the 2005 analysis only produced reliable results up to early to mid-August, yet the whales’ northward migration is thought to extend a few weeks beyond that time. There appears to be considerable merit in revisiting the three data sets and applying as consistent an analysis as possible, in order that more reliable comparisons can be obtained.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the financial support for this project from the Government of Australia, through the Australian Marine Mammal Centre, Tasmania. The collection of the survey data was a huge effort and thanks are due to the land-based survey volunteers (Josh Smith, Wendy Blanchard, Sarah Gardner, Christy

Harrington, Sarah Jossul, Julianne Kucheran, and Melinda Rekhdahl) who endured very basic and trying conditions camping out on Dirk Hartog Island. Special appreciation is due to Verity Steptoe (who co-ordinated much of the aerial survey), for dealing promptly with data queries and questions about the aerial survey component, and to the other aerial observers, including Chris Burton and Kerry Jane Simons. Considerable logistical and other assistance, including transport and loan of equipment, was provided by the Department of Environment and Conservation, Denham – Shark Bay District Manager, Brett Fitzgerald, and Marine Ranger, Wayne Moroney. As for the 2005 survey, Dr Michael Noad, School of Veterinary Sciences, University of Queensland, helped most generously with the loan of equipment and advice. Facilities and administrative and other assistance continued to be provided to JLB at the Western Australian Museum through the courtesy of the Trustees, the Executive Director, and the Head, Science and Culture. Charles Paxton kindly provided the R-project he used for the 2005 analysis, which assisted in comparisons between the two surveys. SLH would like to thank Mark Bravington for most helpful discussions and advice, particularly with estimation of variance.

REFERENCES

Bannister, J.L. 1964. Australian whaling 1963, catch results and research. *CSIRO Division of Fisheroes and Oceanography Reports* 38: 13pp.

Bannister, J.L., Kirkwood, G P and Wayte, S E, 1991. Increase in humpback whales off Western Australia. *Reports of the International Whaling Commission*, 41: 461-465.

Bannister, J.L. 1995. Report on aerial survey and photoidentification of humpback whales off Western Australia, 1994. Report to the *Australian Nature Conservation Agency (unpublished).* 17pp. [Available from Dept of Environment and Heritage, PO Box 787, Canberra, Australia]

Bannister, J.L. and Hedley, S.L. 2001. Southern hemisphere group IV humpback whales: their status from recent aerial survey. *Memoirs of the Queensland Museum.* 47: 587-598.

Bannister, J.L. *et al.* 2009 Report on aerial and land-based surveys of humpback whales off Western Australia, 2008. Report to the *Australian Marine Mammal Centre* (unpublished). April 2009.

Barlow, J., Oliver, C.W., Jackson, T.D. and Taylor, B.L. 1988. Harbor porpoise, *Phocoena phocoena* abundance estimation, for California, Oregon, and Washington: II. Aerial surveys. *Fishery Bulletin* 86(3): 433-444.

Bravington, M.V., Hedley, S.L. and Wood, S.N. (in prep.) Saddlepoint approximations for Poisson and binomial data with overdispersion induced by spatial random fields, with application to clustered line transect data.

Buckland, S.T., Breiwick, J.M. Cattanach, K.L. andLaake, J.L. 1993. Estimated population size of the California gray whale. *Marine Mammal Science.* 9: 235-249.

Dunlop, R. 2008. Land-based survey. Report on land-based survey of humpback whales off Western Australia (unpublished). 21pp.

Hedley, S.L. and Buckland, S.T. 2004. Spatial models for line transect sampling. *Journal of Agricultural, Biological and Environmental* Statistics 9: 181-199.

Kim, Y.J. and Gu, C. 2004. Smoothing spline Gaussian regression: more scalable computation via efficient approximation. *Journal of the Royal Statistical Society, Series B,* 66:337-356.

Marques, F.F.C. and Buckland, S.T. 2003. Incorporating covariates into standard line transect analyses. *Biometrics,* 59, 924-935.

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

Paxton, C.G.M., Hedley, S.L. and Bannister, J.L. (in press) Group IV Humpback whales: their status from aerial and land-based surveys off Western Australia, 2005. *Journal of Cetacean Research and Management, Special Issue on Humpback Whales.*

Pike, D.G., Gunnlaugsson, T. and Víkingsson, G.A. 2008. T-NASS Icelandic aerial survey: survey report and a preliminary abundance estimate for minke whales. Paper SC/60/PFI12 presented to the Scientific Committee of the International Whaling Commission, June 2008. (unpublished). 29pp

R Development Core Team (2008) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

Thomas, L., Laake, J.L., Strindberg, S., Marques, F.F.C., Buckland, S.T., Borchers, D.L., Anderson, D.R., Burnham, K.P., Hedley, S.L., Pollard, J.H., Bishop, J.R.B. and Marques, T.A. 2006. Distance 5.0. Release 2. Research Unit for Wildlife Population Assessment, University of St. Andrews, UK. http://www.ruwpa.st-and.ac.uk/distance/

Wood, S.N. 2006. Generalized Additive Models: An Introduction with R. Boca Raton: Chapman and Hall-CRC. 391pp.

Wood, S.N. 2008. Fast stable direct fitting and smoothness selection for generalized additive models. JRSS Series B. 70: 495-518.

D. Zwillinger (ed.), 2002. Standard Mathematical Tables and Formulae, 31st Edition, CRC, Boca Raton, 910 pp.

APPENDIX: VARIANCE PROPAGATION IN LINE TRANSECT SPATIAL MODELS

A detection function, $g(y; \pi)$, is estimated from some line transect data, y, using some parameter estimates, $\hat{\pi}_y$. The results are used to compute effective strip width (and hence log(effective area)) along the tracklines, and then log(effective area) is included in the spatial model of number of sightings per segment as an offset. Integration over the fitted surface gives total abundance, but uncertainty in $\hat{\pi}_y$ needs to be propagated through to the final abundance estimate in order to estimate the variance in the abundance estimate.

Consider the i^{th} stretch of effort: suppose n_i whales were seen, and that the mean location of the segment was (lat_i, lon_i). Denoting effective area by a_i and using a spatial smooth $s(\cdot)$ to describe spatial abundance, we have

$$\begin{aligned}
\log[E(n_i)] &= \log(a_i) + s(lat_i, lon_i) \\
&= l_i(\pi) + s(lat_i, lon_i) \\
&= l_i(\hat{\pi}_y + \delta) + X_i\beta \\
&\approx l_i(\hat{\pi}_y) + \left[\frac{dl_i}{d\pi}\right]_{\pi=\hat{\pi}_y} \cdot \delta + X_i\beta
\end{aligned}$$

where $l_i = \log(a_i)$, δ is defined as $(\hat{\pi}_y - \pi)$, and X is the design matrix associated with the smoother. Now note that $\left[\frac{dl_i}{d\pi}\right]\delta$ and $X_i\beta$ have identical 'shape' – they are both matrices dotted with vectors. The matrix of first derivatives may be thought of as another 'design matrix' and δ as a vector of unknown parameters. The prior distribution of δ has mean 0 and variance $-H_\pi$, where H_π is the Hessian from maximizing the likelihood of the detection function from the line transect data. The form of the prior distribution for β is also known; it is Gaussian with mean 0 and variance θS^{-1}, where S is the penalty matrix and θ is the smoothing parameter(s) (to be estimated). Thus, δ and β play very similar roles, the only difference being that the 'smoothing parameter' for δ is known (it equals one), whereas for β, it needs to be estimated.