

# Group IV humpback whales: their status from aerial and land-based surveys off Western Australia, 2005

CHARLES G.M. PAXTON\*, SHARON L. HEDLEY<sup>+</sup> AND JOHN L. BANNISTER<sup>#</sup>

Contact e-mail: [cgp2@st-andrews.ac.uk](mailto:cgp2@st-andrews.ac.uk)

## ABSTRACT

Single platform aerial line transect and land-based surveys of Southern Hemisphere Group IV humpback whales were undertaken to provide absolute abundance estimates of animals migrating northward along the western Australian coast during June–August 2005. The aerial survey was designed to cover the whole period of northward migration but the resulting estimates from that survey alone could only, at best, provide relative abundance estimates as it was not possible to estimate  $g(0)$ , the detection probability along the trackline, from the data. Owing to logistical constraints, the land-based survey was only possible for a much shorter period (two weeks during the expected peak of the migration in mid-July). This paper proposes three methods that utilise these complementary data in different ways to attempt to obtain absolute abundance estimates. The aerial line transect data were used to estimate relative whale density (for each day), allowing absolute abundance from the land-based survey to be estimated for the short period of its duration. In turn, the land-based survey allowed estimation of  $g(0)$  for the aerial survey. Absolute estimates of abundance for the aerial survey were obtained by combining the  $g(0)$  estimate with the relative density estimates, summing over the appropriate number of days. The most reliable estimate of northward migrating whales passing the land station for the period of the land-based survey only was 4,700 (95% CI 2,700–14,000). The most reliable estimate for the number of whales passing through the aerial survey region for the duration of that survey (55 days from June through to August) was 10,300 (95% CI 6,700–24,500). This is a conservative estimate because the duration of the aerial survey was almost certainly shorter than the period of the migration. Extrapolation beyond the end of this survey was considered unreliable, but abundance from the estimated start of the migration to the end of the survey (87 days from mid-April to August) was estimated to be 12,800 (95% CI 7,500–44,600). The estimated number of whales depends crucially on the assumed migration and period of migration. Results for different migration parameters are also presented. The point estimates of abundance, whilst higher than those from a previous survey in 1999 (when adjusted for survey duration) are not significantly so. The peak of the whales' distribution was found at c.90m water depth.

KEYWORDS: HUMPBACK WHALE; MIGRATION; ABUNDANCE ESTIMATE; SOUTHERN HEMISPHERE; METHODOLOGY; LINE TRANSECT; SURVEY-COMBINED; MODELLING; TRENDS

## INTRODUCTION

Between 1976 and 1994, a series of aerial surveys was flown at intervals off Shark Bay, Western Australia (WA), to provide information on possible increase in numbers in the Southern Hemisphere, Group IV, population of humpback whales (*Megaptera novaeangliae*). That population, summering mainly in the Antarctic between longitudes 90°E and 135°E, and wintering off the coast of western Australia, was severely depleted on two occasions in the twentieth century, 1934–39 and 1949–63 (Chittleborough, 1965). There were also some substantial but smaller catches from the population before 1934. By the end of 1963, when Australian humpback whaling ceased, the population was estimated to have been reduced to fewer than 600 animals (Bannister, 1964) from c. 7,800 in 1951 (estimates from mortality rates and catch per unit effort by Chittleborough, 1965); at that time it was believed that prior to 1934, the population size might have been as high as 10,200 (from sightings in Chittleborough, 1965).

Following increasing reports of humpback whale sightings in winter off the western Australian coast in the early-mid 1970s, aerial surveys of animals during their northward migration were undertaken from Carnarvon WA (24°52'S, 113°38'E) in an area off Shark Bay where aerial spotter and other data from whaling operations were available for 1963, the last year of humpback whaling.

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 4,000–5,000 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 at least three years would be required between surveys, leading to a proposed further survey in 1997. Given funding constraints, that survey took place in 1999.

The aim of the 1999 survey was to provide a direct estimate of absolute abundance. It thus differed from its predecessors where only a relative index was required. It was planned to cover as much of the northern migration period as possible, with flights every other day over a two month period, from 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 8,200–13,600 (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. Only northward-migrating whales were of interest in that their migration was the one

\* Research Unit for Wildlife Population Assessment, Centre for Research into Environmental and Ecological Modelling, University of St Andrews, The Observatory, Buchanan Gardens, St Andrews, Fife, KY16 9LZ, UK.

<sup>+</sup> The Schoolhouse, Denhead, St Andrews, Fife, KY16 8PA, UK.

<sup>#</sup> c/o The Western Australian Museum, Locked Bag 49, Welshpool DC, WA 6986, Australia.

most completely covered by the planned survey period. This paper details the analysis of data from that further survey, which took place in June–August 2005 with the aim of estimating the size of the northward migrating population of Group IV humpback whales.

## METHODS

### Overview

Estimating whale abundance from stocks such as the eastern Pacific gray whale (*Eschrichtius robustus*) and the east Australian humpback whale, which typically migrate close to the coastline, has been conventionally conducted using land-based (dual-platform) migration counts (Buckland *et al.*, 1993; Noad *et al.*, 2011). For logistical reasons, a similar land-based survey of the western Australian humpback whale stock has not been possible, but funding was obtained to investigate the feasibility of augmenting an aerial survey (to cover as much as possible of the migratory period) with a short land-based survey at the peak of the migratory period. Thus, in 2005, an aerial survey off Western Australia was planned to cover the period from 20 June–20 August (as in 1999), and to follow the same track design (see Bannister and Hedley, 2001), but to be augmented by a land-based component originally planned for three weeks to coincide with the expected peak of the whales' migration. Data from this component of the survey were intended to be used to 'calibrate' the relative density estimates from the aerial survey, so as to gain a better understanding of absolute density. In this paper, three different calibration methods are examined (all of which utilise data from the two components of the survey) and discuss the relative strengths and weaknesses of each method are discussed.

Method A uses the land-based data to estimate the passing rate of humpback whales past the point of the station, and hence to provide an estimate of abundance for the period of the land-based survey. The aerial survey data are used to estimate a relative density surface, which is used to estimate the true local distribution of whales close to the land-based station (since this distribution cannot be assumed to be uniform). The aerial data are also used to extrapolate westward, beyond the visual range of the land-based observers.

Method B directly compares the estimated relative densities from the aerial and land-based surveys (in the same area), to provide a simple calibration of the aerial survey.

Method C uses the aerial survey data to estimate a conventional line transect estimate of surfacing whales (using data from the designed east-westerly oriented transects only). A correction factor to account for whales being unavailable for detection because they were diving is then applied to obtain an estimate of whale density. Surfacing and diving rates were obtained for this correction factor from focal follows made from the land-based station. Method C was applied for comparison with earlier surveys for which no land-based data (but some surfacing and dive time behaviour) were available.

### Field methods

As in the 1999 survey, all aerial survey work was undertaken from a high-wing, twin-engine *Partenavia* P68B aircraft, under charter from Tropicair Services Pty Ltd, flying out of Carnarvon. The survey was originally planned to cover the area surveyed in previous years, approximately 80 n.miles  $\times$  30 n.miles immediately west of Bernier, Dorre and Dirk Hartog Islands on the western boundary of Shark Bay, i.e.

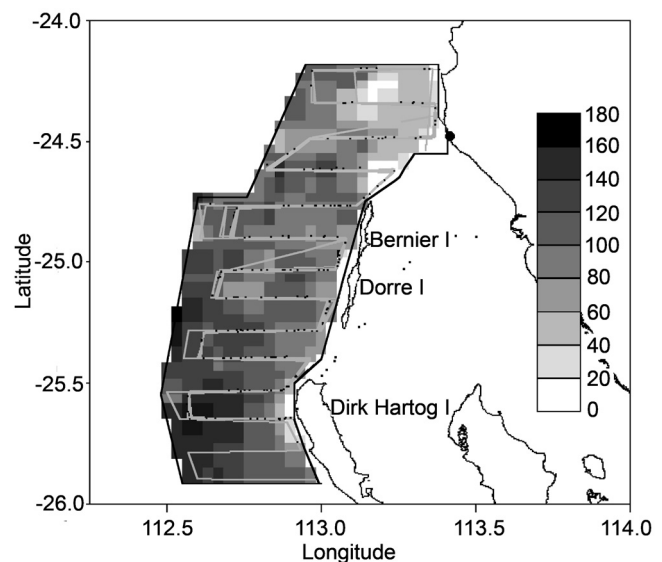


Fig. 1. Aerial tracklines over the entire survey period (grey lines), with locations of associated sightings (small black circles) of NM moving humpbacks at a 2km resolution. Land survey observation point is given by the larger circle situated on the coastline at about 24.5°S. Coloured area gives formal survey region with on effort sightings. Shading indicates depth in metres.

between 112°30'–113°10'E and 24°46'–26°09'S, with transect lines between seven and eight n.miles apart (Fig 1). The land-based station was originally intended to be located on the west coast of Dirk Hartog Island for three weeks, but logistical difficulties meant that an alternative site (at Quobba, 24°29'S, 113°25'E, shown as a large circle in Fig. 1) had to be used. As a result, survey effort was shifted correspondingly northwards, with four legs off the southern part of Dirk Hartog Island being transferred to cover the area north of Bernier Island/Carnarvon, to include the Quobba area. On two occasions, the flight path was extended to the west opposite Bernier Island, in an attempt to check on the distribution of migrating animals beyond the normal search area. A GPS and onboard computer system were available to plot waypoints and record sightings details. Two observers, one on each side of the aircraft used a clinometer (industry standard *Suunto* PM-5/360PC) and an angle board to measure declination and horizontal angles to sightings. The aerial survey took place between 22 June (day 173) and 15 August 2005 (day 227).

Additionally, owing to logistical constraints the land-based survey could only be conducted for a reduced period, namely the two week period from 9–22 July (days 190–203). The observation platform itself was located 1.22km inland. Surveying was carried out for up to 9 hours (8:00 to 17:00) each day with four observers on watch per hour, and was terminated if sea state reached 5 (wind speed 20–25 knots) or in the event of persistent rain. One observer continually scanned the south (190° to 260°) alternating between binoculars and the naked eye, a second did the same to the north (260° to 340°). The third observer operated a theodolite for bearings and declinations. The fourth observer operated the program 'Cyclopes' (see below; the program was developed by Eric Kniest, University of Newcastle, NSW), scanning the entire area when not recording data. A total of six observers was rotated on an hourly basis and none observed for longer than three hours without a break.

For accurate pod tracking, the theodolite was connected to a lap-top running 'Cyclopes'. Whales within a pod were

located with the theodolite and the horizontal and vertical angles recorded. These were plotted on a map as icons linked to a data box, giving accurate distances and bearings of the pod from the observer station. The behaviour, pod composition, migration direction and other comments were filled in by the ‘Cyclopes’ operator. Pods were continually tracked as much as possible using the theodolite to eliminate the possibility of double counting pods. Other observer observations, not fixed by the theodolite, could be filled in as additional observations. Of special interest were the number of animals moving abeam, that is crossing an imaginary line of latitude extending from the land-based survey station, the speed of northern migration and dive time (the last two by reference to measurements made of a focal group of 26 pods). Weather was also recorded on an hourly basis, to include sea state, wind speed (knots) and direction, haze, cloud cover and visibility.

Further details of the survey design and methods are provided in Bannister *et al.* (2006).

**Analysis methods**

The methods described here were applied in an attempt to utilise the combined survey approach to correct for known biases. Conventional estimates from the aerial survey alone can only result in relative abundance estimates since  $g(0)$  is typically less than one. Certain trackline detection (i.e.  $g(0) = 1$ ) is an unrealistic assumption due to ‘perception bias’ (caused by a lack of detection of all whales at the surface) and ‘availability bias’ (whales are unavailable to be detected because they are submerged) as discussed *inter alia* by Marsh and Sinclair (1989). Conventional estimates from the land-based survey alone would likely suffer less from availability bias (since the fixed nature of and wide field of view from the land-observation point means that it is reasonable to assume that almost all whales passing the point would surface in view at least once) but a lack of conspicuousness of a surfacing whale or simply its distance from the observation point may preclude its detection, so perception bias (at least) should probably still be accounted for. However, because only single platform land-based data are available, we must assume that detectability is certain at some distance; the distance selected is zero perpendicular distance (i.e. at the shoreline). The combined operation allowed the three methods to be used to estimate the Group IV population and to correct for the above biases. All three methods utilised the results of the aerial survey so we began by using the aerial survey data to estimate a relative density surface (or ‘spatial model’) of whale pods in the survey region using the following modified version of the ‘count model’ (Hedley and Buckland, 2004; Hedley *et al.*, 1999), followed by the detailed descriptions of Methods A, B and C.

*Aerial analysis*

Transects covered during the aerial survey are divided into small ‘segments’, such that the sighting and geographic conditions do not change considerably within a segment. Unlike the single-stage count model noted above, a two-stage modelling approach is used for this survey, as implemented by Borchers *et al.* (1998). This approach may be considered when there is such a large proportion of segments with zero density that a single-stage analysis (with its associated distributional assumptions relating the variance to the mean) is clearly inappropriate. In the two-stage approach, the presence or absence of whale pods in a segment is first modelled using a logistic regression, and then

the estimated number of whales in the non-zero segments is modelled separately, using quasi-likelihood estimation with variance proportional to the mean (see below for a discussion of the potential biases). Thus in the first stage the Bernoulli data were modelled as

$$E\left(\frac{p_i}{1-p_i}\right) = \exp\left(\theta_0 + \sum_k q_k(z_{ik})\right), \quad i = 1, \dots, T$$

where  $p_i$  is the probability that a whale is present in the  $i^{\text{th}}$  segment,  $\theta_0$  is the intercept parameter, the  $q_k$  are smooth functions of the  $k$  spatial covariates,  $z$ , and  $T$  is the total number of segments. The response variable for the non-zero data model is the estimated density of individuals based on the estimated number of individuals in each segment,  $\hat{N}_i$ , calculated using an estimator similar to the Horvitz-Thompson estimator (Horvitz and Thompson, 1952), as follows:

$$\hat{N}_i = \sum_{j=1}^{n_i} \frac{s_{ij}}{\int \hat{g}_{ij}(y, v) \pi(y) dy}, \quad i = 1, \dots, T$$

where, for each segment  $i$  containing at least one sighting,  $\int \hat{g}_{ij}(y, v) \pi(y) dy$  is the estimated probability of detection (from a line transect analysis, see ‘results’) of the  $j^{\text{th}}$  detected pod,  $n_i$  is the number of detected pods in the segment and  $s_{ij}$  is the size of the  $j^{\text{th}}$  pod. By assumption,  $\pi(y)$ , the probability density function of perpendicular distances, is uniform.

Having obtained the estimated number of individuals in each segment, the estimated density,  $\hat{D}_i$ , is simply given by  $\hat{N}_i/a_i$  where  $a_i$  is the area (twice the truncation width,  $w$ , multiplied by the length) of segment  $i$ . Non-zero pod density  $\hat{D}_i$  (across segments with sightings only) was then modelled as a function of the  $k$  spatially referenced covariates,  $z$ , using a quasipoisson GAM with a log link function:

$$E(\hat{D}_i | \hat{D}_i > 0) = \exp\left(\beta_0 + \sum_k r_k(z_{ik})\right), \quad i = 1, \dots, T$$

where  $\beta_0$  is the intercept parameter and the  $r_k$  are smooth functions of the  $k$  spatial covariates. For this method, only relative densities from the aerial survey data are required, so for estimation purposes, it can be assumed that  $g(0)$  for the aerial survey is 1, and the procedure is valid so long as  $g(0)$  does not vary spatially across the survey region. The area of interest in this case, however, is the region close to the land-based station; the spatial model provides an estimate (albeit a biased one) of the variation in true density in that region.

Sightings were allocated to these segments by reference to their time of observation. In addition to day (Dayofyear), the following spatially-referenced covariates were available for consideration for inclusion in the two-stage model: longitude (Lon); latitude (Lat); water depth (Depth)<sup>1</sup>; distance from eastern boundary of the survey (Disteastbound). Water depths were allocated to effort segments by finding the closest point in the bathymetry data to the midpoints of the effort segments using great circle distances. The bathymetry within the survey region is illustrated in Fig. 1. All covariates were considered for inclusion in the model as 1D smooths (thin plate regression splines) of the untransformed covariate values. In addition, 2D smooths (thin-plate splines, Wood, 2003) of Lat and Lon (suitably transformed) were considered for inclusion into the GAM. 2D and 3D smooths incorporating Dayofyear were also considered. Initially a maximum of six degrees of freedom (seven knots) was allowed in the selection of 1D smooths and up to 13 degrees of freedom (14 knots) were allowed in the case of 2D smooths, thus allowing moderate

<sup>1</sup> Water depths were obtained from ETOPO2 2 minute resolution relief data available from NOAA (<http://www.ngdc.noaa.gov/mgg/global/seltopo.html>).



flexibility but reducing the likelihood of spurious fitting of unnecessarily complicated functions. In addition, the cost associated with fitting each degree of freedom was increased to 1.4 to minimise the risk of overfitting (Kim and Gu, 2004; Wood, 2006). Generalised cross validation (GCV) implemented in the *mgcv* package (Wood, 2001) in *R* was used for covariate selection, augmented with diagnostic plots, using the principles described in Wood (2001) to minimise the GCV with the extra criteria that a term must explain an additional 4% of the overall deviance or variance given other variables in the model and be associated with a significant probability of less than 5%. Owing to gaps in search effort along transects, effort could not always be split into segments of precisely 10km. Therefore, the area of each segment varied slightly around the mean of 100km<sup>2</sup> (= 10km × 2 × 5km) and so the models were weighted by segment area.

#### Method A: Adjusted land-based estimation

Sightings and catch data off WA suggest that northward-migrating whales are concentrated near the shore (Bannister and Hedley, 2001; Dawbin, 1966), with fewer animals seen further offshore. Thus, it is not reasonable to assume a uniform distribution of true abeam (or estimated-abeam) distances from the land-based station. In theory, however, distance sampling techniques can still be applied; the assumption that perpendicular distances are uniformly distributed needs to be modified to one that more closely represents the true distribution of whales from the shore. The program *Distance* (Thomas *et al.*, 2002) can then be used to fit a model to the abeam distances (with no monotonicity constraints). The resultant curve was a composite function,  $f(x)$ , representing the variation in detectability with distance as well as variation in density. If it is then assumed that detectability is certain at zero distance, then this function can be rescaled to intercept the relative density curve ( $d(x)$ ) at the shore. The composite, relative density and (unknown and inestimable) detection functions are related as follows:

$$f(x) = \frac{g(x)d(x)}{\int_0^w g(x)d(x)dx}$$

$$\therefore f(x) \propto d(x)g(x)$$

where  $w$  is the distance at which abeam distances are truncated. In fact, this is a slight simplification: in practice, the spatial model incorporates survey day as a factor, so that the estimated relative whale density can be estimated on a daily basis (to tally more directly with the daily land-based counts). The correction for absolute proportion of whales missed by the land station is estimated as

$$c_\kappa = \frac{\int_0^w g_\kappa(x)d_\kappa(x)dx}{\int_0^w d_\kappa(x)dx}$$

where the  $\kappa$  subscript denotes day. Note that the correction corresponds to the ratio of the areas under the two curves of Fig. 2. The corrected number of pods ( $P_\kappa$ ) passing abeam per day within the visual region of the land-based observers (i.e. in a region truncated at  $w$  from the shore and excluding the land in this case) is then given by:

$$P_\kappa = c_\kappa \times r_\kappa \times 24$$

where  $r_\kappa$  is the number of pods per hour seen from the land-based station (within  $w$  of the shore) on day  $\kappa$ .

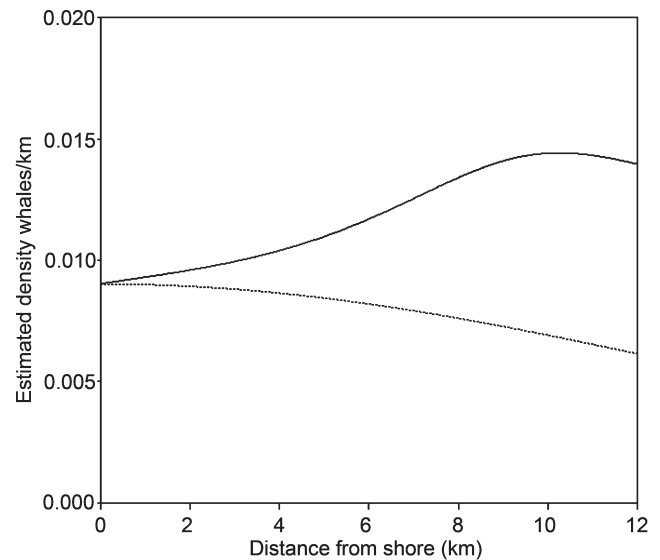


Fig. 2. The estimated availability function (estimated whale density; solid line) and land-based composite function (dashed line) for NM whales on day 190. The area under the solid line is proportional to the total number of pods at that latitude. The area under the dotted line is proportional to the total number of pods detected. The difference between these two areas represents the proportion of pods not seen and allows correction of the land-based estimates. Note: The predictor for the availability function (Depth) was smoothed to produce a smooth curve in the response.

This estimate was adjusted to account for whales passing beyond the visual range of the land-based observers (from the shore to 12km offshore) by considering the relative number of pods from the aerial survey within 12km of the shore to those in the region extending from the shore to the western edge of the survey region – see below).

#### Method B: Calibration of aerial survey

Allowing for the time taken for pods to pass through the survey area, daily estimates of relative abundance may be obtained from the predicted relative density surfaces (obtained as described in Method A). By reference to the days when the land-based survey and aerial survey overlapped, estimates of  $g(0)$  are obtained from the quotients of the daily land-based estimates of number of pods to those from the aerial survey (and thus a mean  $g(0)$  of the aerial survey could be crudely estimated), correcting for both availability and perception bias.

#### Method C: Availability at the surface from focal pod observations

An alternative distinct method of estimating  $g(0)$  is to construct an estimate from direct observations of diving pods made during the land-based survey. Bannister and Hedley (2001) estimated  $g(0)$  for aerial surveys of this population in 1999 using the following equation (from Barlow *et al.*, 1988):

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

where  $s$  = average time spent above the surface,  $t$  = window of time during which an animal is within the visual range of an observer and  $d$  is the average time below the surface. For the present analysis, we derived a modification of this estimator as:

$$P(\text{visible}) = \frac{s}{s+d} + \frac{t[1-s/(s+d)]}{d_{\max}} \quad \text{for } t > 0$$

where  $d_{max}$  is the maximum (observed) dive time. This assumes that if the whales are observed for a period of time greater than  $d_{max}$  then the animal will be observed if within the visual range of the observer. Using estimates of surfacing and diving times from focal studies on 26 pods conducted during the land-based survey, this estimator is used to provide estimates for the aerial survey which account for availability bias. Method C is applied in this paper to provide a comparison of results from other analyses (Bannister and Hedley, 2001) which also did not adjust for perception bias.

All three methods described above assume that the aerial survey covered the entire east-west width of the migration path. This may not be true (see below) in which case the resulting estimates represent an underestimate of the total number of whales.

**Variance estimation**

Variance was estimated using the non-parametric bootstrap (Efron and Tibshirani, 1993). Appropriate sampling units (assumed independent) are transect legs for the aerial survey and days for the land survey. This should ensure good bootstrap coverage for both surveys (samples are generated with replacement). Variance due to the migration speed of the whales was incorporated by sampling with replacement from the observed distribution of pod swimming speeds. Variance in surface availability for the  $g(0)$  estimate derived from Method C is not readily incorporated – a weakness of the method.

**RESULTS**

**Aerial data**

Details of the aerial data are summarised in Table 1. Initially, all data available from north of 25.9°S outside the bay, i.e. to the west of the islands and up to 24.2°S (see Figs 11a and b, results sets 1–6;9–14, Table 2) were used in the analysis. The total usable aerial trackline length was 7,500km. There were 417 initially usable sightings of pods (mean size ± standard error, 1.81±0.04) within the entire aerial survey region; their locations are shown in Fig. 1. One aim of the survey was estimating the relative abundance of northward migrating (NM) humpback whales, and yet a ‘northward migrating’ whale is not easily identifiable in the field. Aerial observers recorded swimming direction; those recorded as swimming north ( $n = 178$ ) formed the NM data subset. Whales of unknown swimming direction ( $n = 133$ ) were randomly allocated a migration direction in proportion to the sightings of known direction on the same day. Those that were allocated a northerly swimming direction were added to the NM sightings and this formed a possible NM dataset (hereafter termed ‘NM+’). The remaining sightings were either classified as swimming south ( $n = 96$ ) or milling ( $n = 10$ ). Examination of the distribution of perpendicular sighting distances suggested that some sightings close to the trackline were missed (Fig. 3), a feature of aerial surveys

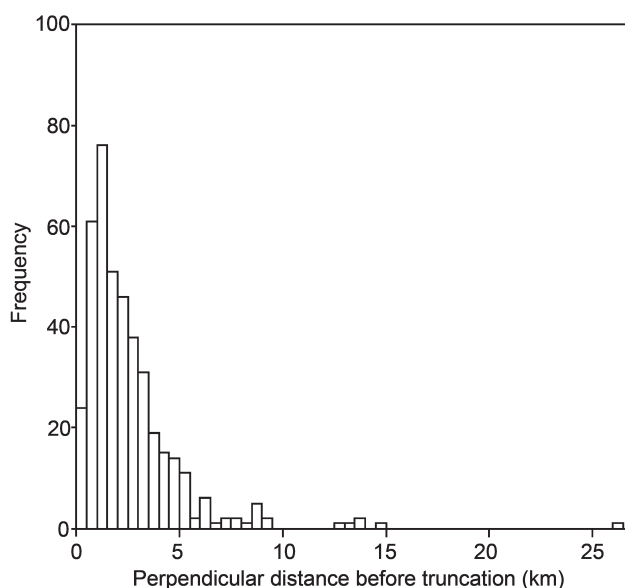


Fig. 3. Frequency distribution of horizontal perpendicular distances from the aerial survey of all on transect, in survey zone, sightings of humpback whales ( $n = 417$ ).

when, as for this survey, the aircraft is not fitted with bubble windows, so the data were left-truncated at 1km. The distance data also exhibited a long tail, with one sighting reported at over 25km which perhaps seems implausible. This sighting, along with others seen at over 6km from the trackline, were excluded from the analysis, resulting in a total of 303 sightings (of any swimming direction). Subsequent analysis was also performed on a data set from 25.9°S to 24.5°S ( $n = 224$ , see Figs 12a and b, results sets 7–8, Table 2). The former are here referred to as the ‘main’ data set (and grid) and the latter the ‘southern’ data set (and grid). Use of the southern data subset potentially reduced the chance of a milling pod to be misclassified as a NM pod.

**Estimation of relative density from the aerial survey data (for Methods A, B and C)**

First it was necessary to estimate the probability of detection of each pod sighted on the aerial survey. A multiple covariate distance sampling (MCDS) approach was adopted (Marques, 2001; Marques and Buckland, 2003) in which the effects of covariates (in addition to perpendicular distance) can be incorporated into the detection function via the scale parameter. Available covariates and factors were: Observer; Beaufort sea state; Cloud cover; Pod size; and Sightability.

A stepwise forward selection procedure (starting with a model containing perpendicular distance only) was used to decide which covariates to include in the model, as indicated by Akaike’s Information Criterion (AIC). To determine the aerial detection function all available data in the survey region were considered regardless of recorded swimming direction.

Table 1  
Sample sizes used in the survey. All available sightings were used in the detection function ( $n=303$ ).

Data	Sightings (mean ± SE)	Pod sizes (mean ± SE)	Total number of segments	Segment area km <sup>2</sup> (mean ± SE)
Whole data set certain northern migrating whales (NM)	137	2.06 ± 0.08	749	100.70 ± 0.52
Whole probable northern migrating whales (NM+)	192	1.91 ± 0.06	749	100.70 ± 0.52
Southern data set certain northern migrating whales (NM)	116	2.07 ± 0.09	587	101.00 ± 0.59
Southern data set probable northern migrating whales (NM+)	155	1.94 ± 0.07	587	101.00 ± 0.59

Table 2

Estimates of abundance of western Australian humpback whales by various methods. All methods assumed a mean pod size of 2.051. Hourly migration was taken to be 5.35km/h. All aerial survey based models error estimates include a migration component unless otherwise stated. All aerial predictions based on a grid between  $-24.2^{\circ}$  Lat and  $-25.9^{\circ}$  Lat approximately apart from results sets 7–8.

Results set	Data set	Method	Detection/availability estimates and 95% confidence intervals	Date of time period [Days]	No. of days	Estimates and 95% confidence interval	Relevant figures
1	NM	A. Passing whales on land based survey. Relative abundance from aerial survey	Assumed land based $g(0) = 1$ . Proportion of pods missed within the 12.5km truncation limit = 0.341. Proportion of aerial survey area at latitude covered by land based survey = 0.177.	9 Jul.–20 Jul.; 22 Jul. [190–201; 203]	12	5,500 (3,000–25,900)	Figs 7, 8 and 9
2	NM+	A. Passing whales on land based survey. Relative abundance from aerial survey	Assumed land based $g(0) = 1$ . Proportion of pods missed within the 12.5km truncation limit = 0.342. Proportion of aerial line of latitude covered by land based survey = 0.211.	9 Jul.–20 Jul.; 22 Jul. [190–201; 203]	12	4,600 (2,700–14,000)	Figs 7, 8 and 9
3	NM	Aerial survey with $g(0)$ adjusted by reference to land based survey	Aerial $g(0) = 0.268$ (0.069–0.437)	27 Jun.–15 Aug. [173–227]	55	9,200 (5,700–34,000)	11a
4	NM+	Aerial survey with $g(0)$ adjusted by reference to land based survey	Aerial $g(0) = 0.310$ (0.125–0.492)	27 Jun.–15 Aug. [173–227]	55	10,300 (6,700–24,500)	11b
5	NM	Assumed uniform density from aerial survey (EW transects only) with $g(0)$ taken from focal pod observation	Aerial $g(0) = 0.279$	27 Jun.–15 Aug. [173–227]	55	7,200 (5,000–10,400)	
6	NM+	Assumed uniform density from aerial survey (EW transects only) with $g(0)$ taken from focal pod observations	Aerial $g(0) = 0.279$	27 Jun.–15 Aug. [173–227]	55	8,400 (6,100–11,400)	
7	NM	Aerial survey but predictions made across grid and from data to $-24.5$ Lat only	Aerial $g(0) = 0.279$	27 Jun.–15 Aug. [173–227]	55	10,000 (7,100–14,000)	12a
8	NM+	Aerial survey but predictions made across grid and from data to $-24.5$ Lat only	Aerial $g(0) = 0.279$	27 Jun.–15 Aug. [173–227]	55	12,700 (9,100–17,200)	12b
9	NM	Aerial survey with $g(0)$ adjusted by reference to land based survey	Aerial $g(0) = 0.268$ (0.073–0.437)	20 Apr.–15 Aug. [140–227]	88	9,400 (6,000–36,000)	
10	NM+	Aerial survey with $g(0)$ adjusted by reference to land based survey	Aerial $g(0) = 0.310$ (0.126–0.499)	20 Apr.–15 Aug. [140–227]	88	12,800 (7,500–44,600)	
11	NM	Aerial survey with $g(0)$ adjusted by reference to land based survey	Aerial $g(0) = 0.268$ (0.070–0.443)	11 Jun.–1 Sep. [162–242]	81	15,400 (8,500–56,500)	
12	NM+	Aerial survey with $g(0)$ adjusted by reference to land based survey	Aerial $g(0) = 0.310$ (0.100–0.500)	11 Jun.–1 Sep. [162–242]	81	15,500 (9,500–47,500)	
13	NM	Aerial survey with $g(0)$ adjusted by reference to land based survey	Aerial $g(0) = 0.268$ (0.081–0.439)	1 Jun.–8 Sep. [152–252]	101	22,500 (10,000–72,200)	
14	NM+	Aerial survey with $g(0)$ adjusted by reference to land based survey	Aerial $g(0) = 0.310$ (0.114–0.496)	1 Jun.–8 Sep. [152–252]	101	19,400 (10,800–59,700)	

Further, to test whether detectability differed with recorded direction of movement, Direction was also considered as a factor in an analysis of a subset of the data where direction was known. All model selection was performed in Distance (v4.0, Thomas *et al.*, 2002), then the final selected model (a half-normal model with Observer fitted as a factor – see Fig. 4) was re-fitted using a set of unpublished customized functions in the statistical package *R* v.2.2.1. (Ihaku and Gentleman, 1996). This facilitated estimation of variance within *R* (see below). There was no evidence of detectability of whales varying with known direction of movement, nor was there evidence that pod size influenced detectability presumably owing to the lack of substantial variation in pod size. Therefore all sightings of whale pods within the truncation distance and survey area were considered to generate detection probabilities for subsequent analysis although only potentially northward migrating (NM+ or NM) pod sightings were used to estimate density. The same model was also fitted to the southern data set.

As described earlier, the number of sightings of NM pods was used to estimate a density surface for NM whales. Transects covered during the survey were divided into segments. Choice of segment length is a compromise between maximising the ratio of non-zero to zero segments, maintaining environmental resolution and giving some measure of spatial independence. We selected a segment length of 10km after some preliminary investigation of these aspects.

The final selected logistic model for presence-absence included two 1D smooths: Dayofyear (with 5 degrees of freedom) and Depth (with 5 degrees of freedom). Estimated probabilities of presence for each day are plotted against water depth in Fig. 5. The model explained about 8% of the deviance on the linear predictor scale in the case of NM whales and 5% in the case of NM+ whales. This was a binary model, so the explained deviances do not necessarily provide information about the agreement between the observations and the fitted probabilities (Collett, 1991; McCullagh and

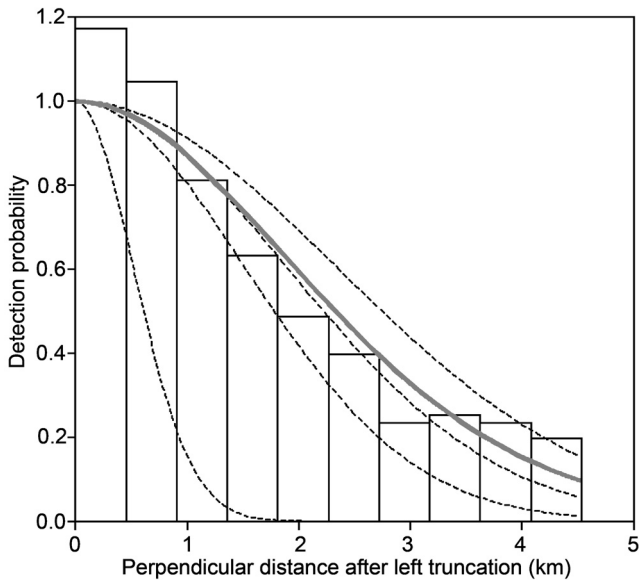


Fig. 4. Mean aerial detection function based on 303 humpback whale sightings between 1 and 6km apart. Dotted lines indicate the detection curves for individual observers. The solid line is the mean detection probability.

Nedler, 1989). For both datasets, extrapolation from the model suggested that there could be whales to the west of the survey area. The probability of humpback presence peaked at *ca* 90m water depth. The second stage of the density surface modelling used those segments of trackline in which sightings were made ( $n = 112$ , in the case of NM whales;  $n = 156$  in the case of NM+ whales). In the case of the NM whales, the estimated along-trackline densities ( $\hat{D}$ ) were between 0 and 0.165 whales/km<sup>2</sup> (mean: 0.008 whales/km<sup>2</sup>). In the case of the NM+ whales, the estimated densities were between 0 and 0.182 whales/km<sup>2</sup> (mean: 0.010 whales/km<sup>2</sup>). The estimated pod densities were not spatially correlated except possibly at very small distances (by inspection of a semi-variogram). Using a logarithmic link function and assuming that the variance of the observations was proportional to their mean, the best-fit model for NM

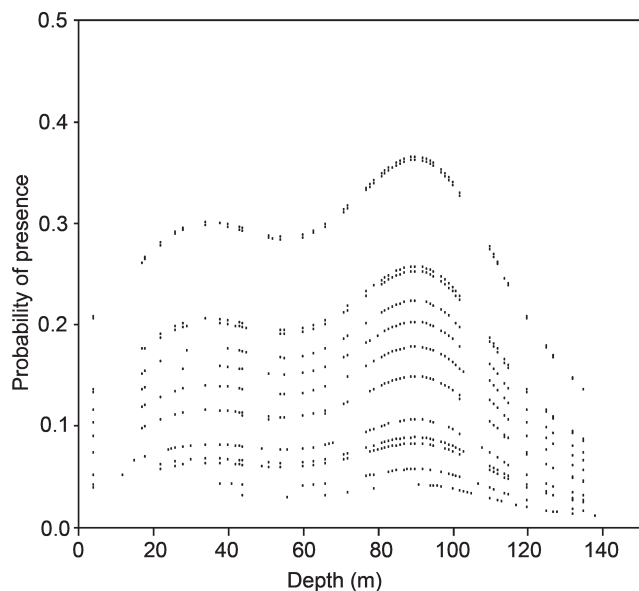


Fig. 5. Probability of presence of whale pods on different days (the separate curves) against depth.

whales included just one covariate – Depth (with 4 degrees of freedom) – and explained just 4% of the deviance. (Depth itself was non-significant both in the case of NM whales and NM+ whales, but according to other model selection criteria, such as GCV as outlined above, it was retained in the model.) Thus while water depth predicted pod presence it did not predict numbers of pods (although it was retained in the model for estimation).

Relative densities within the survey region were estimated for NM pods and for NM+ pods, based on the selected models for NM pods (Figs 11a and b), as products of the two predicted surfaces from the two-stage modelling. The region is bounded, therefore estimates should be considered as conservative because the survey does not appear to have covered the full longitudinal range of the whales' migration path. The corresponding estimated density surfaces within the smaller southern grid are shown in Figs 12a and b.

Prediction from the two-stage model throughout the survey region and across a range of days from day 100 (10 April) to day 230 (22 August) suggested that the peak day of the northern migration was day 191 (10 July). Extrapolation of the trend to the early part of the year suggested that there was negligible probability of the migration starting before April and that numbers only really started increasing in May. However such extrapolation assumes that the trend in pod numbers can be extended into the past and future. The model also predicted a second pulse in the northern migration from day 220 (12 August; see Fig. 6) for which there was only weak evidence from the data. We speculate that this is due to the poor temporal coverage achieved, resulting in little data to 'tie down' the tail of the migration period. An alternative approach would have been to assume a more traditional unimodal migration curve, and restrict the flexibility afforded to the GAMs accordingly.

**Land-based survey analysis (Method A)**

Having estimated the relative density surface within the survey region from the aerial survey data, it was assumed that this represented the shape of the true distribution of

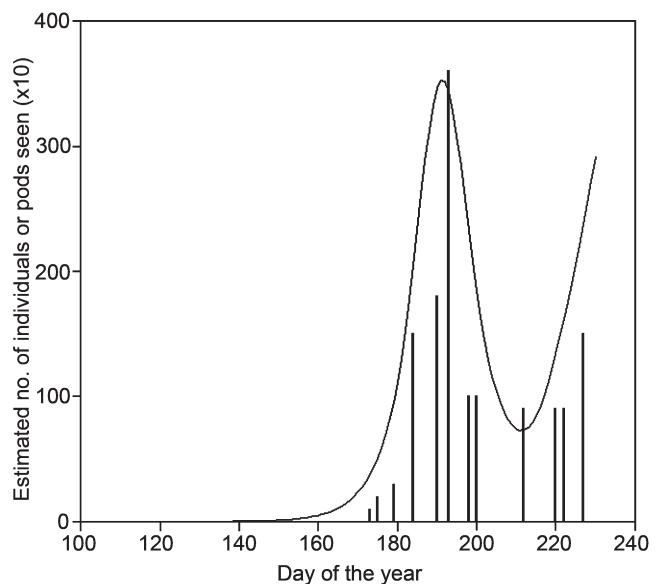


Fig. 6. Estimated number of available detected NM humpback individuals present assuming an instantaneous aerial survey of the shaded region in Fig. 2 from 10 April (day 100) to 22 August (day 230) (curve). Pod sightings per day ( $\times 10$ ) are given as the bars. N.B. the ( $\times 10$ ) is arbitrary to allow the data and the model predictions to be compared.



whales in the search region of the land-based observers. The land-based survey data were then used to estimate the ‘composite’ function of ‘abeam’ distance,  $x$ , from the shore; a function whose shape was determined by the decreasing detectability of pods with distance as well as varying underlying whale density. The land-based survey was undertaken from 9–22 July (days 190–203), but only one hour of effort was possible on day 202 (during which time no whales were seen); this day’s effort was subsequently excluded from the analysis. Of 235 land-based sightings (mean pod size 1.62), 99 were classified as NM whales (mean pod size 1.79) of which 41 were actually seen abeam. Data were left-truncated at 1.22km (as the land station was located 1.22km inland) and right-truncated at 13.22km (to make estimation of the composite function more robust). The frequency distribution of distances to all 99 NM pods is given in Fig. 7; 93 of these were between 1.22 and 13.22km. The perpendicular distances were binned into 2km classes and the composite function of perpendicular distance from shore was fitted in *Distance*, allowing for the possibility of adjustment terms and with no monotonicity constraints. In addition, potential explanatory variables were Beaufort sea state (considered as continuous and as a factor), pod size and visibility (both continuous). The selected model was a simple function of perpendicular distance (Fig. 8). As detailed above, the hourly passing rates of NM pods (depicted in Fig. 9) were then adjusted to give estimates of the daily number of NM pods passing the observation point within 12km of the shore. The peaked distribution in time seen in the aerial survey data was not evident in these data, apart from a peak at day 199 (8 days after the aerial survey peak) presumably because of the limited time frame and paucity of sightings.

When this survey was planned, the land-based station was originally to be sited on the west coast of Dirk Hartog Island – a location where previous aerial surveys had indicated a relatively longitudinally-narrow migratory corridor (Bannister and Hedley, 2001), in which it was expected that the great majority of sightings would pass within the visual search region of the land-based observers. In the event, the

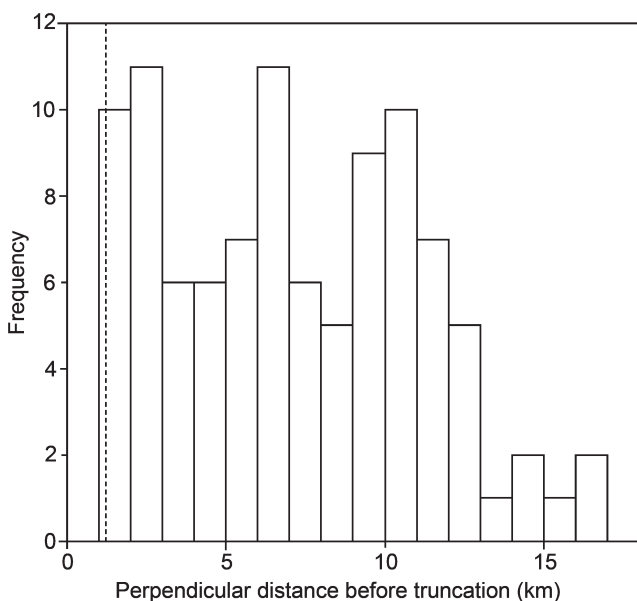


Fig. 7. Frequency distribution of horizontal distances to NM pods crossing either abeam of the land survey location or otherwise minimum distance to the shore ( $n = 99$ ). Dashed line shows the shore line. Note: The lowest bar is artificially low as part of its bin region was in fact on shore.

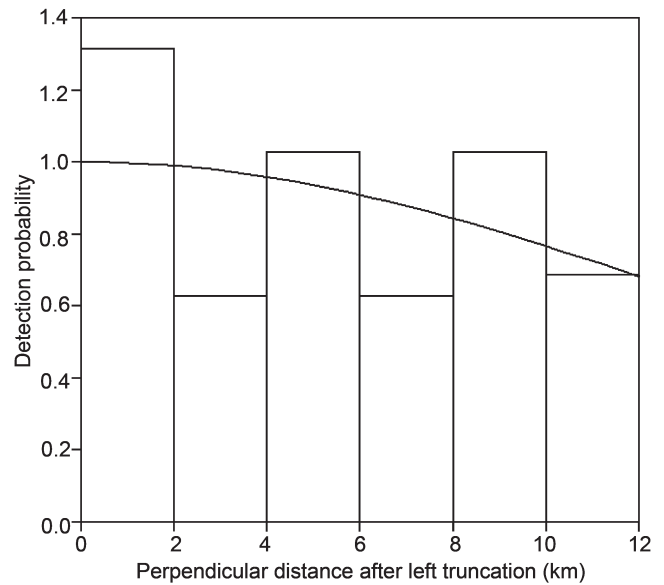


Fig. 8. Land-based abeam composite function for northward migrating whales ( $n = 93$ ).

land-based station had to be located at Quobba. However, even off Dirk Hartog Island further south (Fig. 1), the aerial survey data indicated that whales were distributed beyond the visual range of the land-based observers. Here, therefore, we use the ratio of estimated relative density curves at the latitude of the land station to estimate the total number of whales (to the western edge of the survey area, at 58.7km from the shore) passing the observation point during the period of the land-based survey as follows:

$$\frac{\int_0^{12} \hat{d}(x) dx}{\int_0^{58.7} \hat{d}(x) dx}$$

The estimated NM mean pod size (2.05) from the aerial survey was then multiplied by the reciprocal of the above to obtain the estimated numbers of whales to the western edge of the survey area. Estimates were then summed over days

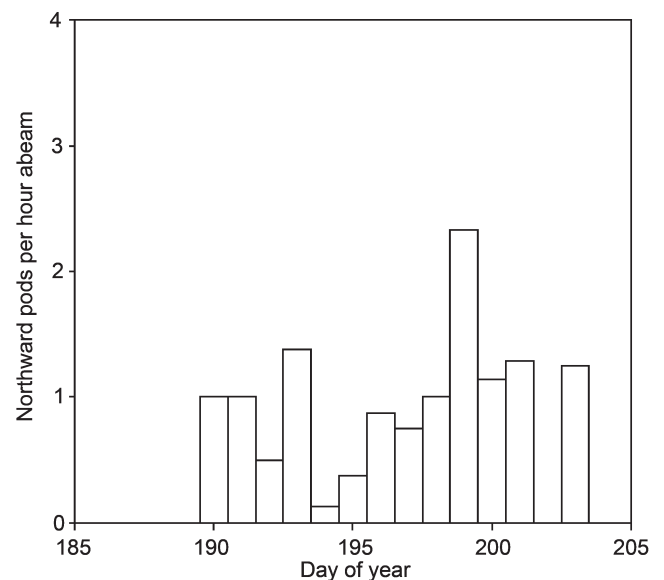


Fig. 9. Per hour passing rate for the land-based survey.



190–201 and day 203, to obtain an estimate of 5,500 NM whales (95% CI 3,000–25,900) during that period. The comparable estimate for NM+ whales was 4,600 NM (95% CI 2,700–14,000). Whilst the difference between these point estimates is not significant, it is somewhat counterintuitive (a consequence of the different GAMs in the two-stage density surface modelling). This suggests that alternative models for the density surfaces (perhaps using either zero-inflated models or the spatial models being developed by Bravington *et al.* (2008)) should be investigated in future.

#### Land-based behavioural analyses (for Methods B and C)

Estimated migration speeds (for Method B), diving and surfacing rates (for Method C), were derived by data from following 26 focal pods during the land-based survey. The mean observed migration speed was  $5.35\text{kmh}^{-1}$  (SE = 0.58). The 26 focal pods were observed for an average of 31 minutes each and mean  $s$  and mean  $(s + d)$  were calculated from the observed time submerged and the total time observed.

From the 26 individuals pods investigated, the maximum dive time recorded was 1,173 seconds. Assuming that, if for a given pod the basic instantaneous probability of being observed is  $s/(s + d) = 0.1723$  and that it is definitely observed if watched for a period of time equal to the maximum dive length recorded (1,173 seconds), and if the function connecting those two points is treated as linear, then as a crude approximation:

$$P(\text{visible}) = 0.1723 + 0.0007t \quad \text{for } 0 < t \leq 1,173 \text{ seconds}$$

The window of time observed was calculated as per Bannister and Hedley (2001) from the frequency distribution of distances ahead, abeam and aft of the observers. The vast majority of sightings (99%) were between 2.6km aft and 7.4km forward (Fig. 10) implying a total window of opportunity of 10km. With the mean aircraft speed of 127.5 knots (=  $236\text{kmh}^{-1}$ ) then this distance would be travelled in 153 seconds implying, from equation (2), that the probability of observing a pod at the surface on the aerial survey is 0.279. This assumes that the observed maximum of 1,173 seconds is not substantially different from the maximum possible dive time during migration.

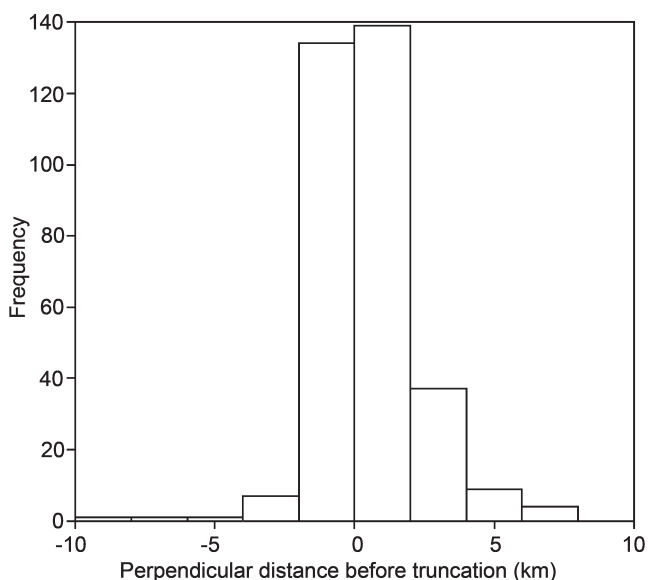


Fig. 10. Frequency of afore, abeam and aft horizontal distances to observed pods in the aerial survey.

#### Calibration of aerial survey (Method B)

For each relevant day, the predicted instantaneous number of detected pods was obtained as the product of the predictions of the logistic regression and the non-zero density regression as described above. These numbers were then modified by a correction factor that reflected the time taken for a pod to pass through the survey area (latitudinal range  $1.735^\circ$ , or 192.7km). For example, if the NM whales travelled at mean speed  $5.35\text{kmh}^{-1}$  (the mean speed from the land based survey), then the correction was

$$(5.35 \times 24) / 192.7 = 0.665$$

Given the daily estimate of available whales from the aerial survey and the estimate of whales from the land-based survey, an estimate of  $g(0)$  for the aerial survey is given by the quotient of the estimated number of pods from the aerial survey to the estimated number of pods from the land-based survey. For NM whales, was about 0.29, whilst for NM+ whales it rose to about 0.33. The higher in the latter case could be reflecting the ‘detection’ of unspecified-direction whales at the surface.

#### Abundance estimates

A wide selection of abundance estimates is presented in Table 1. Variants were:

- (1) Aerial sightings dataset used: NM or NM+.
- (2) Method: land; aerial with  $\hat{g}(0)$  from land-aerial calibration; aerial with  $\hat{g}(0)$  from observational data; aerial with assumed equal to 1; restricted aerial (E–W transects only and a uniform density model) with  $\hat{g}(0)$  from observational data; aerial with  $\hat{g}(0)$  from observational data but prediction only over a subset of the region.
- (3) Period of days over which estimation is made: 190–201 and 203 (land method only); 173–227 (period of 2005 aerial survey); 162–242 (80 day migration period); 152–252 (100 day migration period); 140–227 (87 day period covering the start of the migration and the period of the 2005 aerial survey).

Summaries of the most important results from Table 2 are given below.

#### Results sets 1–2 (Method A. Land-based survey period: days 190–201, 203)

The total estimate of NM whales is 5,500 and for NM+ whales is 4,600. The land-based estimates assume that all whales in the survey area pass through the latitude of the land survey within a band at  $24.48^\circ\text{S}$  of width 58.7km from  $113.41^\circ\text{E}$  to  $112.85^\circ\text{E}$ . These estimates are lower than all aerial estimates (except those which assumed  $g(0) = 1$  – see below) but they are estimated over the shorter time frame of the land-based survey only.

#### Results sets 3–8 (Methods A and B. Aerial survey period: days 173–227)

Applying the  $g(0)$  corrections derived from the land-based survey results in a total whale abundance of 9,200 NM whales for the duration of the aerial survey (results set 3). For NM+ whales the corresponding estimate is 10,300 (results set 4). Fig. 11 gives a density surface for (a) NM and (b) NM+ whales using the estimated mean migration speed from the focal pod study and the land-based survey-derived  $g(0)$  estimates. Results sets 5 and 6 display the results of the analysis of the east-west transects only assuming a uniform density, as an equivalent to a conventional distance analysis.

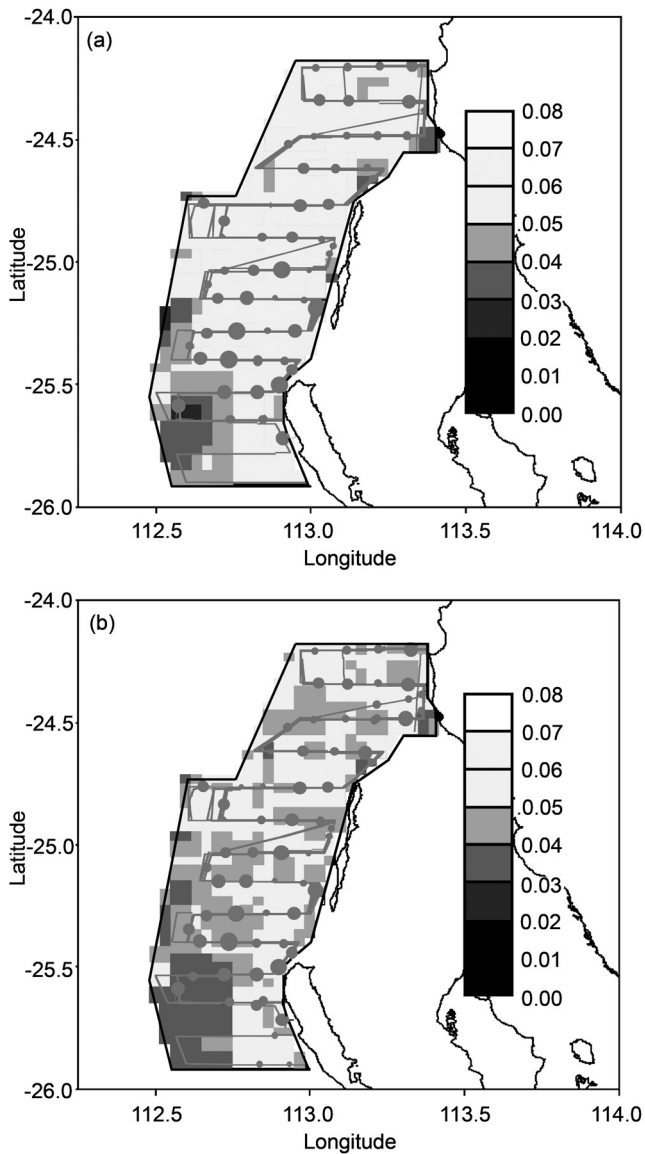


Fig. 11. Estimated densities for northward humpback whales (individuals per km<sup>2</sup>) 2005 corrected for availability by reference to the land based survey (result sets 3 and 4) illustrated for the estimated peak of the migration at day 191. Grey lines are the transect lines and the area of the grey circles is proportional to the summed Horvitz-Thompson like estimate of the number of whales per segment. The black circle on the coast at 24.5° latitude is the location of the land based survey. Grey line shows boundary of survey region (a) NM whales, (b) NM+ northward whales.

The  $g(0)$  estimate of 0.279 (from focal pod observations) fell within the range of that generated by the land-based survey. As a consequence the abundance estimates nested within the range of the estimates found in results sets 3 and 4. The analyses based on the southern data only and with  $g(0)$  again estimated from focal pod observations as 0.279 (results sets 7–8, Fig. 12) generated similar results to the other estimates presented for this survey period (except those which assumed  $g(0) = 1$ ).

*Results sets 9–11 (Estimated start of migration to end of aerial survey: days 140–227)*

Fig. 6 suggests reasonable confidence in extrapolating beyond the range of the (aerial survey) data prior to its commencement, but the presence of the apparent second pulse at the end of the survey means that the results from any extrapolation after the end of the survey should be treated extremely cautiously. In this result set, therefore, abundance

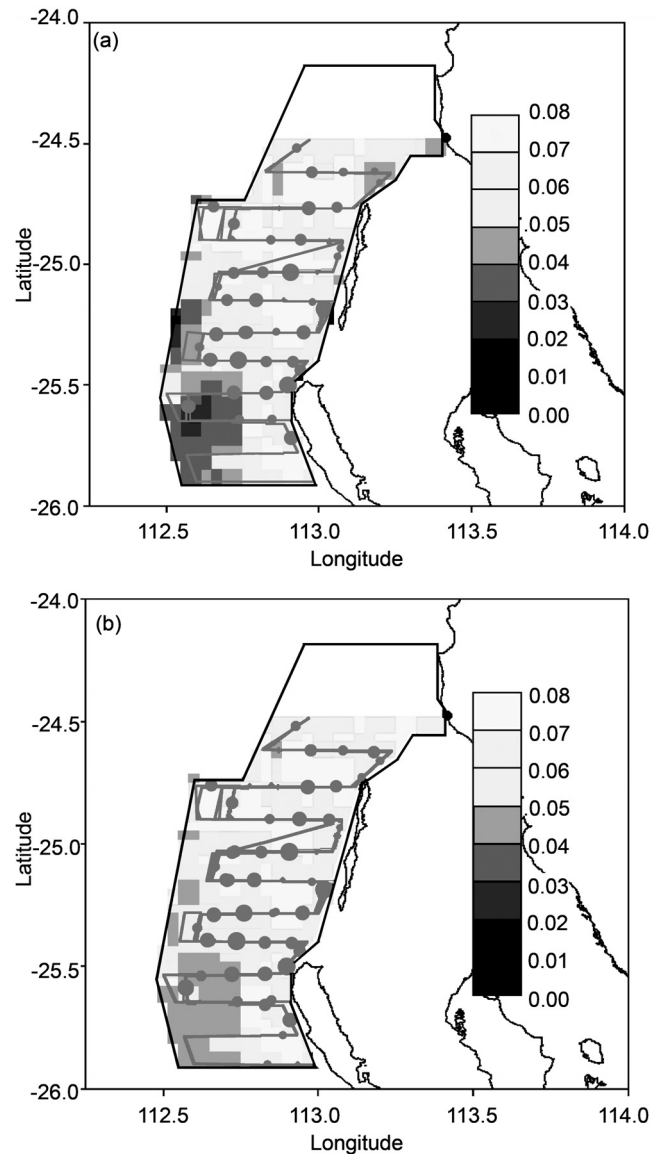


Fig. 12. Estimated densities for northward humpback whales (individuals per km<sup>2</sup>) in 2005 assuming  $g(0) = 0.279$ , for the region south of 24.5°S. Result sets 7 and 8, illustrated for the estimated peak of the migration at day 191. Grey lines are the transect lines and the area of the grey circles is proportional to the number seen. The black circle on the coast at 24.5° latitude is the location of the land based survey. (a) NM whales, (b) NM+ whales.

was calculated from the estimated start of the migration but only until the end of the aerial survey. Applying the  $g(0)$  corrections derived from the land-based survey results in a total whale abundance of 9,400 NM individuals. For NM+ whales the corresponding estimate is 12,800.

## DISCUSSION

This analysis represents a first attempt to apply a density surface to model distributional heterogeneity in a southern humpback population, although a similar method has been applied to North Atlantic humpback whales (Paxton *et al.*, 2007). There is a large amount of unexplained variance. Nonetheless all the available aerial data analyses suggest a population of about 10,000 individuals, albeit with substantial variance around that figure. However a major influence on the size of the estimate was its temporal range.

In the absence of a full scale land-based (independent observer) survey for the full period of the humpback whale migration, we have attempted to provide estimates of absolute abundance from a series of aerial surveys intended to cover the entire time period and a land-based survey that covered about 10–15% of the whole migration period. The aerial survey analysis provided ‘snapshot’ estimates of the relative density surfaces of a moving population of whales, a large proportion of which remain unavailable for detection to the aerial observers because they are diving whilst in the visual range of the observers. The initial analysis problem was thus twofold: to estimate the numbers that were below the surface and to convert snapshot estimates into estimates allowing for whale movement.

### Surface availability

Two methods were used to assess the availability of animals at the surface in the aerial survey. Arguably the two methods are estimating different things. The estimate based on observed dive times is just for surface availability. The estimate based on the land survey is for surface availability plus under-detection on the trackline (in fact in this case at 1 km from the trackline because of left-truncation). However under-detection on the trackline is possibly negligible given the similarity of the estimates. Using the land-based survey to estimate  $g(0)$  has the considerable advantage that its uncertainty can be readily included in the final estimate. Therefore,  $c.0.34$  of northward migrating animals are detected at the surface on the trackline. The higher end of this range for  $g(0)$  is for analyses which incorporated some unknown-direction whales re-classified as northward whales, and as would be expected, this is higher than for analyses which used only whales with northward direction recorded on the survey (the lower end of the range).

### Potential problems with the methods

The methods as implemented above could have a number of biases, as listed below.

- (1) The two-stage zero-inflated modelling of pod densities could lead to an under-estimation of whale pod numbers as many zeros, if the pods are at low densities, are not true zeros but ‘presences’ with a low non-zero density. Further the choice of the quasipoisson distribution for the non-zero data was problematic. Finally the resolution of the model was crude to compare with the land based composite function.
- (2) The consideration of NM-only data could lead to an underestimation of northward pod numbers if a disproportionate number of unknown direction pods are northward whales.
- (3) The random allocation of probable northward status to unknown direction whales could lead to bias if the unknown whale pods did not represent a random selection of northward and southward whales.
- (4) The trend in presence of pods with Depth as well as with Dayofyear suggests that only a subset of whales in space and time were enumerated, therefore there is a risk of underestimating abundance.
- (5) The land-based calibration assumes that the distribution of whales relative to the shore is proportionately constant every day. In this analysis, no evidence was found of an interaction of Dayofyear with Depth (i.e. that distribution changed with depth but at a different rate each day). If this were to be the case then this would have to be incorporated into the density surface estimation. Further

problems could arise if there was a systematic shift in response to transient environmental variables such as sea surface temperature.

- (6) It is assumed all submerged whales crossing abeam at distance 0 are seen at some stage.
- (7) In the estimates that are based on the calibration of the aerial surface with the land based survey, the swimming speeds are used twice introducing an element of circularity.

### Comparison with the 1999 survey: are humpback whale numbers increasing?

Bannister and Hedley (2001) offered a variety of point estimates derived from the 1999 survey for two periods of 80 days and 100 days respectively (11 June–1 September and 1 June–8 September) during the migration season. Applying an approximate estimate of  $g(0)$  of 0.3 (from all analyses here) to their best unbiased surface individual estimates of 3,365 for 80 days and 3,441 for 100 days and crudely multiplying up their confidence intervals, gives new abundance estimates for 1999 of 11,200 (95% CI 9,000–13,900) and 11,500 (95% CI 9,200–14,300) respectively.

Using the generated estimates for whales from analysis set 4 (i.e. with estimates of  $g(0)$  from focal pod observations) but extended to cover the estimation range of the 1999 survey gave point estimates (and confidence intervals) for NM+ whales of 15,500 (9,500–47,500) for 80 days (results set 12, Table 1) and 19,400 (10,800–59,700) for 100 days (results set 14, Table 1). The variances are higher in the latter case because the new method includes uncertainty in  $g(0)$  even though presumably more variation in the density surface is explained. Despite the point estimates being higher for 2005 compared to the 1999 estimate using the same surface availability estimate, the estimates are not significantly different. For completeness, the corresponding estimates for NM whales only are also presented in Table 1 (results sets 11 and 13).

### Environmental preferences

The apparent peak in the presence of pods at 90m reflects qualitative observations of humpback whale preferences for shallow water of the coast of Australia. Shallow water preferences in humpback whales have been quantified both in feeding (Paxton *et al.*, 2007) and nursing/breeding periods (Ernst and Rosenbaum, 2003) but less so during migration. Whether this depth preference is an active preference by whales or the outcome of some other mechanism remains to be elucidated.

### Suggestions for future work

Further thought should be given to establishing the boundaries of the whale migration in space and time. Coverage of regions to the west of the current survey region could determine the western boundary of migration. Using a density surface could allow the boundaries to be readily determined even if surveying is patchy. The extrapolation of the temporal trend suggests that migration could continue from May; given the suggestion of a second pulse to the migration (Fig. 6), it is unclear what occurs after the survey period. Intensive overlap of land-based surveys and aerial surveys could enhance the estimation of  $g(0)$  especially if coupled with fine resolution environmental data (e.g. depth) which could allow reliable interpolation of the distribution of animals from the shore. A double platform analysis with two or more independent observers would provide an alternative method for estimating  $g(0)$  in both the aerial and



land-based surveys. The most powerful advantage of an aerial survey is that it allows the distributional properties of the population to be mapped onto environmental features providing an insight into the biology of the animals.

Future work (perhaps using zero-inflated models) could better model the aerial density surface leading to more precise estimates of abundance.

The logistical changes to the land-based survey meant that its location was not ideal, as the whales were often seen milling. Whales passed by the land station far beyond the visual range of the observers, but in principle, that should not prove problematic for this combined analysis because the (relative) density of whales was also estimated from the aerial data. What the analysis did highlight, however, was the need for a reliable and informative model of the density distribution close to the land station. Another consideration for any land based survey is that the available sector for observation at distance 0 effectively encompasses the dive time of the whales. Failure to capture a wide enough window about the abeam line from shore will result in a bias if all north moving pods are not observed at distance 0.

Further investigation is required to investigate the possible 'reasons' for the somewhat peculiar shape of the migration curve (Fig. 6). Had more of the scheduled flights been able to have been completed, a unimodal curve might have been obtained, from which abundance estimates beyond the range of the data could more confidently be predicted. For this analysis, however, we prefer to adopt a cautious approach, preferring the abundance estimates from only the time covered by the aerial survey (or alternatively, those from the start of the migration to the end of the aerial survey), i.e. the estimates from results set 4 (or alternatively set 12) of 10,300 with 95%CI [6,700–24,500] (or 12,800 with 95% CI [7,500–44,600]) although the land-based survey results represent the best estimates for the days of that survey only.

## ACKNOWLEDGEMENTS

The authors would like to thank Tiago Marques and Stephen Buckland for their most helpful discussions. The paper has been greatly improved following the suggestions of two referees. The aerial survey was undertaken under the supervision of Chris Burton (Western Whale Research, Perth) and the land-based operations under Rebecca Dunlop (School of Veterinary Sciences, University of Queensland, Brisbane). Considerable assistance, including the loan of a computer-linked theodolite, was provided by Michael Noad (School of Veterinary Sciences, University of Queensland). JLB thanks the Trustees, Chief Executive Officer and staff of the Western Australian Museum for continued provision of facilities and assistance. The project was funded through the Australian Department of the Environment and Heritage (now the Department for Environment Water, Heritage and the Arts), Canberra, from the Natural Heritage Trust, with particular assistance from Robyn McCulloch.

## REFERENCES

- Bannister, J.L. 1964. Australian whaling 1963. Catch results and research. *Aust. CSIRO Div. Fish. Oceanogr. Rep.* 38: 1–13, plus 7 tables and 4 figs.
- Bannister, J.L. 1995. Report on aerial survey and photoidentification of humpback whales off Western Australia, 1994. 17pp. [Unpublished report to the Australian Nature Conservation Agency (Environment Australia, Canberra)].
- Bannister, J.L., Burton, C.L.K. and Dunlop, R.A. 2006. *Humpback whale aerial survey, Shark Bay area, Western Australia, 2005. Report on work done to 31 March 2006 under the Consultancy Agreement between the Australian Government and the Western Australian Museum signed on 9 March 2005.* Dept. of Environment and Heritage, Canberra, Australia. (Unpublished). [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. *Mem. Queensl. Mus.* 47(2): 587–98.
- Bannister, J.L., Kirkwood, G.P. and Wayte, S.E. 1991. Increase in humpback whales off western Australia. *Rep. int. Whal. Commn* 41: 461–65.
- 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. *Fish. Bull.* 86(3): 433–44.
- Borchers, D.L., Buckland, S.T., Preide, I.J. and Ahmadi, S. 1998. Improving the precision of the daily egg production method using generalized additive models. *Can. J. Fish. Aquat. Sci.* 54: 2727–42.
- Bravington, M., Hedley, S.L., Wood, S.N. and Peel, D. 2008. Estimating Antarctic minke whale abundance from SOWER data, using spatial modelling and point independence. Paper SC/F08/A9 presented to the SOWER Abundance Workshop, 7–10 February 2008, Seattle, USA (unpublished). 19pp. [Paper available from the Office of this Journal].
- Buckland, S.T., Breiwick, J.M., Cattanch, K.L. and Laake, J.L. 1993. Estimated population size of the California gray whale. *Mar. Mammal Sci.* 9(3): 235–49.
- Chittleborough, R.G. 1965. Dynamics of two populations of the humpback whale, *Megaptera novaeangliae* (Borowski). *Aust. J. Mar. Freshwater Res.* 16(1): 33–128.
- Collett, D. 1991. *Modelling Binary Data*. Chapman and Hall, London. 369pp.
- Dawbin, W.H. 1966. The seasonal migratory cycle of humpback whales. pp.145–70. In: Norris, K.S. (eds) *Whales, Dolphins, and Porpoises*. University of California Press, Berkeley and Los Angeles. xv+789pp.
- Efron, B. and Tibshirani, R.J. 1993. *An Introduction to the Bootstrap*. Chapman and Hall, New York. 436pp.
- Ernst, P.J. and Rosenbaum, H.C. 2003. Habitat preference reflects the social organisation of humpback whales (*Megaptera novaeangliae*) on a wintering ground. *J. Zool., London.* 260: 337–45.
- Hedley, S. and Buckland, S.T. 2004. Spatial models for line transect sampling. *J. Agric. Bio. Env. Stat.* 9: 181–99.
- Hedley, S., Buckland, S.T. and Borchers, D.L. 1999. Spatial modelling from line transect data. *J. Cetacean Res. Manage.* 1(3): 255–64.
- Horvitz, D.G. and Thompson, D.J. 1952. A generalisation of sampling without replacement from a finite universe. *J. Am. Stat. Assoc.* 47: 663–85.
- Ihaku, R. and Gentleman, R. 1996. R: a language for data analysis and graphics. *J. Comp. Graph. Structures* 5: 299–314.
- Kim, Y.J. and Gu, C. 2004. Smoothing spline gaussian regression: more scalable computation via efficient approximation. *J. Roy. Stat. Soc. (B)* 66: 337–56.
- Marques, F.F.C. 2001. Estimating wildlife distribution and abundance from line transect surveys conducted from platforms of opportunity, Ph.D. thesis, School of Mathematics and Statistics, University of St Andrews, UK.
- Marques, F.F.C. and Buckland, S.T. 2003. Incorporating covariates into standard line transect analyses. *Biometrics* 59: 924–35.
- Marsh, H. and Sinclair, D.F. 1989. Correcting for visibility bias in strip transect aerial surveys for aquatic fauna. *J. Wildl. Manage.* 53(4): 1017–24.
- McCullagh, P. and Nedler, J.A. 1989. *Generalized Linear Models*. Chapman and Hall, London.
- Noad, M.J., Dunlop, R.A., Paton, D. and Cato, D.H. 2011. Absolute and relative abundance estimates of Australian east coast humpback whales (*Megaptera novaeangliae*). *J. Cetacean Res. Manage. (special issue 3)*: 243–252.
- Paxton, C.G.M., Burt, M.L., Hedley, S.L., Vikingsson, G.A., Gunnlaugsson, T. and Desportes, G. 2007. Using environmental covariates to estimate the density of humpback whales with application to the NASS-95 and NASS-2001 aerial shipboard surveys. *NAMMCO Scientific Publications* 7: 143–60.
- 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. and Pollard, J.H. 2002. *Distance 4.0. Release 1*. Research Unit for Wildlife Population Assessment, University of St Andrews, UK. [Available from: <http://www.ruwpa.st-and.ac.uk/distance/>].
- Wood, S.N. 2001. mgcv: GAMs and generalized ridge regression for R. *R News* 1: 20–25.
- Wood, S.N. 2003. Thin-plate regression splines. *J. Roy. Statist. Soc.* 65: 95–114.
- Wood, S.N. 2006. *Generalized Additive Models: an introduction with R*. Chapman and Hall, Boca Raton, Florida. 391pp.