Mark-recapture analysis of sperm whale (*Physeter macrocephalus*) photo-id data from the Azores (1987-1995)

J.N. MATTHEWS[*], L. STEINER[#] AND J. GORDON[+]

Contact e-mail: whalesong@ifaw.org

ABSTRACT

Population estimates for female or immature male sperm whales (*Physeter macrocephalus*) in a region within the Azores archipelago are given, based on photo-identification studies with mark-recapture analysis. The study area encompassed the Central Group of islands and the island of São Miguel. Data indicate that the recapture rate of animals likely to be females differs from other animals, and this heterogeneity is incorporated into the models. Closed population (Petersen) estimates, using data from within summers, suggest a population of between 300-800 female or immature sperm whales in summer in the study area. Estimates of the population that visits the study area in summer were made using a model selected from the Jolly-Seber family. The open population visiting the study area appeared to vary between about 400-700 between the years 1988-1990, increasing by a factor of three to about 1,600-2,200 between the years 1991-1994. The fraction of whales which are not suitably marked for identification is estimated to be about 12%, so these estimates should be increased by a factor of 1.14. These estimates are reliable if the study area covers the range of a wider population which moves into and out of the study area randomly. The increase in abundance in 1991 is probably due to change in the composition of the population visiting the area. It is not yet possible to clearly define the wider population that the Azores samples are from, nor are migration patterns to and from the area understood. Investigations on a larger spatial scale are needed for a better understanding.

KEYWORDS: SPERM WHALE; MARK-RECAPTURE; ABUNDANCE ESTIMATE; NORTHERN HEMISPHERE; NORTH ATLANTIC

INTRODUCTION

The Azores have long been known as both a breeding and feeding area for sperm whales, and as a whaling region (Clarke, 1956; 1981). Sperm whales were taken at least as far back as 1765. Whaling continued, using traditional methods, until recent times: the last whales were captured in 1987. In the latter part of the 20th century the Azorean catch was one of the largest in the North Atlantic. Now, with the development of whalewatching operations, sperm whales remain an important economic resource to the Islands.

Population trends following the end of whaling are important, as the impact of whaling on the population can far outlast the whaling itself. Whitehead *et al.* (1997) have shown that, off the Galapagos Islands, the population of sperm whales decreased by about 20% each year between 1985-1995. They suggest that the continued decline is the residual impact of the mainland whaling industry, which ended in 1981. How the Azorean sperm whales fared after whaling ceased is not known, as long-term knowledge of the population is quite limited. Catch data are reported by Clarke (1956) and Avila de Melo and Martin (1985), and density estimates were made in 1988 and 1989 using acoustic survey methods by Leaper *et al.* (1992).

Whitehead *et al.* (1997) suggest the overexploitation of mature males as one cause of the population decline off the Galapagos. If this is so, then the Azorean population might be expected to show similar trends. Clarke (1956) writes that, according to Drouet, '… in the sperm whale fishery the Azores were noted for large whales (1861)', and in the Azorean records from the mid-20th century onwards, males formed the major part of catches (Avila de Melo and Martin, 1985). Males were often preferred by whalers due to their larger size, or in some northern areas were simply more available. The lengths of whales caught at several North Atlantic fisheries declined after the 1950s, providing an indication of overexploitation (e.g. Martin, 1981). The decline in lengths in the Azores was less clear-cut than in, for example, Iceland, and there was evidence of an increase in the proportion of large whales caught in the last decade of the hunt (Avila de Melo and Martin, 1985).

The primary interests of this paper are the size and conservation status of the population of sperm whales in and around the Azores, in the context of relatively recent large-scale whaling activity and its current economic importance to the whalewatching industry. The study estimates abundance, in the nine years following the end of whaling, using photo-identification and mark-recapture techniques. This approach has been used to estimate abundance in various long-term studies of cetaceans (e.g. see review in Hammond, 1986). Previous mark-recapture estimates of sperm whale abundance have been made off Kaikoura, New Zealand (Childerhouse *et al.*, 1995), and off the Galapagos Islands (Whitehead, 1990; Whitehead *et al.*, 1992). Hammond (1986) stresses the importance of population definition in such studies. In this study, the 'closed' populations of sperm whales are defined to be those animals in the study area in each study season. The 'open' population of sperm whales is defined to be those animals in the wider area that visit the study area over several seasons.

METHODS

Fieldwork and field methods

Fig. 1 shows the study area and the locations of all encounters with sperm whale groups. Data were collected between May and September inclusive from *Song of the Whale*, a 14m auxiliary powered ketch (between 1987 and 1995) and from *Colomban* an 18m schooner (between 1993 and 1995). No data were collected in 1992. Since these

---

[*] *Song of the Whale Research Team, International Fund for Animal Welfare, 87-90 Albert Embankment, London, SE1 9UD, UK.*
[#] *Whale Watch Azores, 6b South Street, Banbury, OX16 7LF, UK.*
[+] *SMRU, Gatty Marine Laboratory, University of St Andrews, St Andrews, Fife, KY16 8LB, UK.*

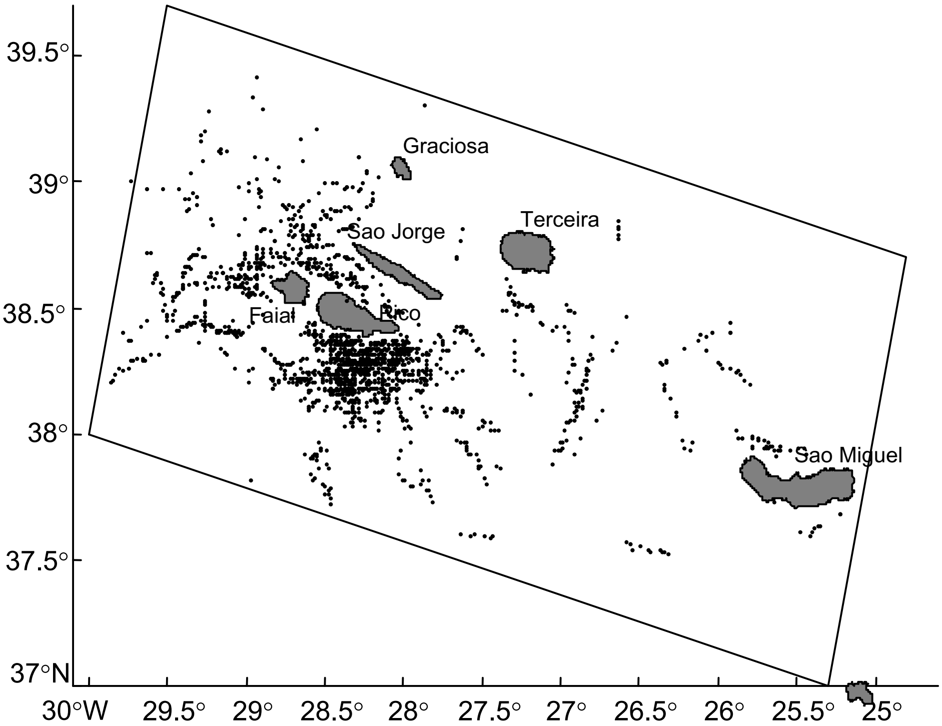

Fig. 1. Study area around the Central Group of islands (Faial, Graciosa, Pico, São Jorge, Terceira) and the island of São Miguel. Points where photo-identifications were made are also shown.

vessels' activities were not dedicated solely to photo-identification, in common with many studies of this type (e.g. Calambokidis *et al.*, 1990; Katona and Beard, 1990) sampling was neither completely random nor systematic. Table 1 shows the length of the field seasons and years of operation of the two vessels.

Table 1

Summary of length of field seasons (days) and number of identified animals for IFAW and WWA.

Year	Length of season (IFAW)	Length of season (WWA)	Total number of identified whales
1987	20	-	21
1988	82	-	151
1989	100	-	105
1990	107	-	94
1991	132	-	159
1992	-	-	-
1993	36	53	97
1994	-	114	93
1995	76	49	141

Individual sperm whales can be reliably identified from various nicks and marks on the trailing edge of the flukes (e.g. Arnbom, 1987; Gordon, 1987). Previous studies have shown that the majority of sperm whales have distinctive flukes suitable for photo-id (of animals with good quality photographs: >91%, Arnbom, 1987; 100%, Childerhouse *et al.*, 1995; 82%, Dufault and Whitehead, 1995). In most cases, these marks appear not to change quickly over periods of years. Childerhouse and Dawson (1996) examined photographs of 56 males seen between one and 56 months apart, all of which were re-identified correctly in a blind experiment. Thirteen showed some change over time but only one showed a 'major' change. Dufault and Whitehead (1995) made comparisons of photographs of flukes of 161 sperm whales encountered between one and eight years apart, 63 of which had changed over time, with only six 'major' changes.

Typically, sperm whales were located using passive acoustic techniques (e.g. Whitehead and Gordon, 1986). In these waters, sperm whales were generally found in scattered groups, the composition of which may reflect long-term associations. Once detected, acoustic methods and sightings were used to locate sperm whales when they came to the surface to breathe. Sperm whales in deep diving (probably feeding) groups typically spend around 10 minutes on the surface between dives of around 40 minutes (Gordon and Steiner, 1992). Such whales typically surface in small groups of 1-3 and move very slowly at the surface, on a constant heading, blowing regularly. To obtain photographs for identification, the research vessel would be carefully and quietly manoeuvred directly behind the whale. If time allowed, the whale's dorsal fin was carefully observed for traces of a callus using 7 × 50 binoculars and, if possible, photographs of the dorsal fin were taken. Sequences of identification photographs were taken (on *Ilford* XP2 film using a *Canon* T90 camera and a 300mm, f4 fixed focal length lens) when the whale raised its flukes above the water to initiate a deep dive. Usually, the research vessel would spend the remainder of a day collecting fluke photographs from a dispersed group. Because of their long dive times, it was rarely possible to collect fluke photographs systematically. On some occasions, a dispersed group might be followed overnight and identification effort would continue during the next day.

Photographs of individuals with sufficiently good marks to allow identification were organised into a catalogue and matched to new photographs by eye. The catalogue was maintained and all matching completed by one of us (LS). The data used here were the time, date, location and identity of each photograph for which identification was possible, and auxiliary information used to classify whales.

Analysis

Population estimation

Over short time periods the assumption of population closure may be adequate, even in an open population. In this study, closed-population estimates were made for each season. Each year's data were divided to form two samples based on time: before 1 August (sample 1) and on or after 1 August (sample 2) and Petersen estimates calculated for each season. In the longer term, processes of birth, death, immigration and emigration are expected. The present data spanned nine years, so 'open' models which incorporate these processes were applied. Using summers as sampling units and a yearly sampling interval gives larger sample sizes and numbers of recaptures (Hammond, 1986).

Open models were selected from the Jolly-Seber (JS) family (Pollock *et al.*, 1990; Schwarz and Arnason, 1996). The fully-parameterised JS model (Jolly, 1965) gives sample-by-sample estimates of abundance, 'survival' (i.e. the proportion of animals that do not emigrate or die) and 'birth' (i.e. the number of immigrants or births). Members of the JS family are modifications of this with constraints placed on the parameter set. For example, survival may be constrained not to vary from sample to sample but to remain constant throughout the study. Using fewer parameters can make the model easier to interpret and may improve precision.

Schwarz and Arnason (1996) provide a formulation of the JS model which overcomes certain technical difficulties, and allows the application of very general models, including modelling by covariates and stratification into multiple groups. In this formulation, JS family models have parameters drawn from the set $\{\phi_1, \ldots, \phi_{K-1}; p_1 \ldots p_K; b_0 \ldots b_{K-1}\}$ where K is the number of samples, ϕ_i the probability of survival from sample i to i+1, p_i the probability of capture at sample i, and b_i the normalised birth rate for sample i. Where appropriate, the population may be further stratified into G sub-groups, giving a total parameter set of G(3K-1).

Models are described by a vector (ϕ_q, p_r, b_s) (following Lebreton *et al.*, 1992; Schwarz and Arnason, 1996), where the subscripts q,r,s are one of :

g	capture constant over time but varies with group (G parameters);
t	capture varies with time but not with group (K parameters);
*g*t*	capture varies with time and group (GxK parameters);
none	capture is a constant across all groups and times (1 parameter).

An example: the model (ϕ_g, p, b_t) has survival rate constant in time but differing between groups, a constant capture probability and time-varying birth rates that do not differ by group. These models were applied using the program POPAN5 (Arnason *et al.*, 1998). Adjustments to survival rates were also made for the year (1992) with missing data.

Care must be taken with the specification of these general models to keep track of 'non-identifiable' parameters (Arnason *et al.*, 1998). In the full JS model, within each sub-group the parameter combinations (ϕ_{k-1}, p_k), (b_0, p_1) and (ϕ_1, p_1, b_1) are non-identifiable, i.e. cannot be estimated separately. As a consequence, some parameters of interest are not estimable in this study (specifically, ϕ_1, ϕ_{k-1}, B_1, B_{k-1} , N_1, N_k).

Stratification

In mark-recapture studies, heterogeneity of capture probability causes biased estimates of population size. Stratification of the population into sub-groups with (more) homogeneous capture probabilities partially addresses this problem. The population was stratified into those individuals which were thought to be probably females, because they had been noted as having dorsal fin calluses or escorting or observed suckling calves on at least one occasion ('probable females'); mature males, based on size or head proportions ('males'); and the remainder ('other'). Observations of mature males, which probably migrate from the Azores to more northern waters, such as those off Iceland (Martin, 1982), were rare and excluded from the population analyses. The population estimates made here are thus for the female and immature male component of the sperm whale population only.

The results indicated that probable females are more likely to be recaptured than others within seasons. Because of these different recapture rates, Petersen estimates for the stratified population were made assuming that the probability of capture is equal between the two sub-groups in the first sample, but different in the second sample. Let n_1, n_2 be the number of captures in the first and second samples, and m_2 the number of marks in the second sample. The pooled abundance estimate is (Seber, 1982):

$$\hat{N} = \frac{(n_1+1)(n_2+1)}{(m_2+1)} - 1$$

For a sub-group of abundance X, with number of captures in sample 1 of n_{1x}

$$\hat{X} = \frac{n_{1X}.\hat{N}}{n}$$

In the remainder of the paper 'F' and 'O' are used as subscripts to denote probable females and others respectively.

Goodness of fit

Pollock *et al.* (1985; 1990) describe goodness-of-fit tests for the Jolly-Seber model, based on χ^2 statistics from pooled contingency tables. These are now familiar as TEST 2 and TEST 3 (Lebreton *et al.*, 1992). The statistic χ^2_{gof} = (TEST 2 + TEST 3) provides a test of the goodness-of-fit of the full Jolly-Seber model. In forming the statistic, individual tables containing cells with expected values less than two are omitted (Pollock *et al.*, 1990). This value is also used in variance estimation, as described below.

Variance estimation

With the probable exception of older males, sperm whales are sociable animals. As a consequence of their gregariousness, encounters with individual animals tend not to be independent of one another. However, these social groups are not wholly stable. Social groups themselves may associate for periods of days, and there is also a certain amount of fission and fusion of these groups. Whitehead *et al.* (1991) have proposed a simple model where, among females and immatures, there is a mixture of relationships, some 'constant' (lasting for years) and some 'casual' (lasting for days). This dynamic and gregarious sociality causes difficulties in analysis, as neither individuals nor groups can be considered as completely independent units.

Anderson *et al.* (1994) consider the statistical effects of schooling: 'Members of such populations can be expected to have a positive correlation among individuals; such dependence causes overdispersion … [T]he estimators of model parameters often remain unbiased …, but the model-based, theoretical variances are underestimated'. Where there is moderate dispersion (a common occurrence in capture-recapture data) they recommend the estimation of a 'variance inflation factor', c. Each empirical variance estimate $\hat{\mathrm{var}}(\theta_i)$ of parameter θ_i is increased to $c.\hat{var}(\theta_i)$. In practice, a single average $c = \chi^2_{gof}/df$ is often reasonable for all variance estimates (Anderson *et al.*, 1994).

Model selection

Parsimonious models can be selected using likelihood-based criteria. With overdispersed data, as in the present study, the model with minimum (modified) Akaike Information Criterion (QAIC) is selected, as recommended by Anderson *et al.* (1994), defined as:

$$QAIC = \frac{-2.\log[L_{\max}(\hat{\theta})]}{\hat{c}} + 2.p$$

where $L_{\max}$ is the maximum-likelihood, c is the variance inflation factor, $\hat{\theta} = \{\hat{\theta}_1, \hat{\theta}_2, ..., \hat{\theta}_p\}$, the vector of parameter estimates, and p is the number of parameters. For ease of reading, details of the model selection process are given in the Results section.

RESULTS

A total of 2,355 identifications were made from photographs, representing 762 individual sperm whales. Of the identified whales, 45 were mature males, 285 were probable females and 432 were others. Identifications could not be made from about 12% of attempted captures with suitable photographs, because the whales were not clearly marked. Sampling effort varied considerably from year to year. The length of field seasons, number of active days within field seasons, and priority that could be given to photo-identification varied, and in later years (1993 and 1995) there were two boats at work (Table 1). The probability of capture of whales is therefore expected to vary from year to year.

Within each season, the majority of individuals were seen on only one day, but some individuals were resighted over the season. There was a tendency to resight probable females over a longer period within a season. Fig. 2 shows histograms of the number of days between first and last sighting by sub-group for animals seen on more than one day. The proportion of resightings within a season that were over one day apart was 31% for probable females and 12% for others. Probable females were also more likely to be resighted between seasons than others. Fig. 3 shows histograms of the number of years between first and last sighting for individuals resighted between seasons by sub-group. The proportion of individuals sighted over more than one season was 18% for females and 6% for others. Among animals resighted between seasons, probable females tend to be seen over a longer period. These patterns justify the stratification of the data into two sub-groups in both closed and open population models.

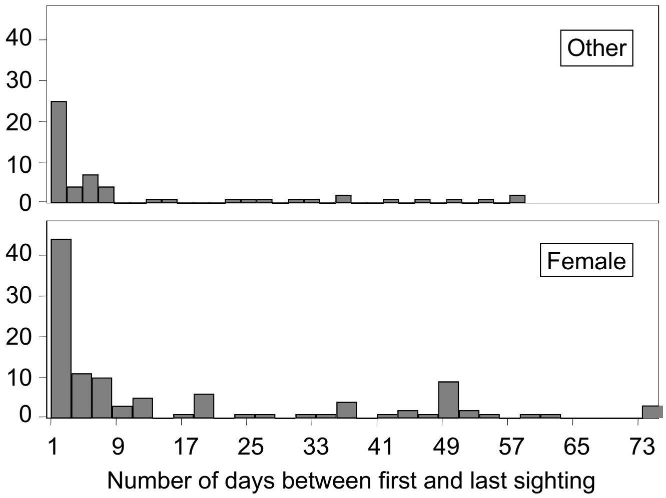

Fig. 2. Histogram of number of days between first and last sighting within each season, for each individual seen on more than one day, by sub-group.

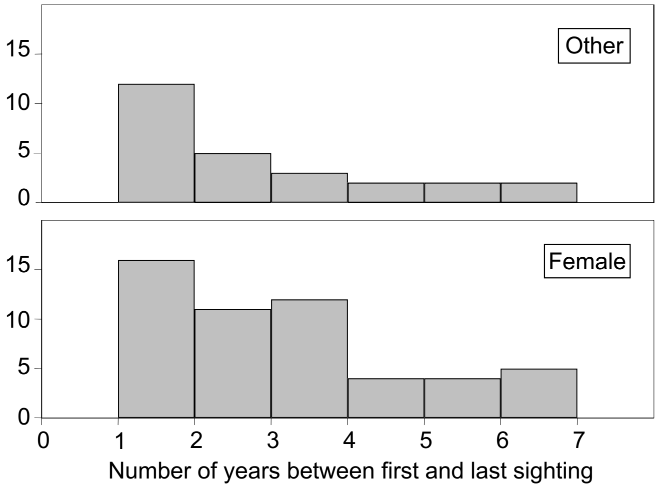

Fig. 3. Histogram of number of years between first and last sighting for each individual seen in more than one season, by sub-group.

The number of years seen are shown in Table 2, broken down by sub-group. There were no between-year resightings of mature males. The probable females sub-group is further split in this table into those animals which had been seen nursing and those which had not (these latter classed as female on the basis of callus information only). There is no obvious difference in relative between-year resighting frequency, so the pooling as probable females was retained.

Within-season (closed) population analysis

Table 3 shows within-season Petersen estimates for different years. The pooled estimates are fairly consistent from year to year, and suggest a population of about 300-800 animals in

Table 2

Frequency of between-season captures. Counts are for the three subgroups by number of years seen. Probable females are also split into two subgroups (females nursing or not nursing).

		Number of years seen			
		1	2	3	4
Other		406	23	3	0
Probable females	Nursing females	190	29	10	0
	Not nursing females (but callus present)	43	6	6	1
	Both	233	35	16	1
Mature males		45	0	0	0

Table 3

Summary of within-season Petersen analysis. Number of captures and recaptures of females and others in the first and second periods of each season are shown. Petersen estimates of population size are shown for the subgroups and for the two combined, with standard deviations in parentheses. See Methods for further details of notation. No data were collected in 1992.

	Observed data					Estimates		
Season	n_{1O}	n_{1F}	n_1	n_2	m_2	N_O	N_F	N
1988	50	55	105	49	6	360 (184)	396 (201)	756 (240)
1989	30	27	57	54	10	152 (57)	137 (52)	289 (67)
1990	27	15	42	49	3	345 (306)	192 (174)	537 (220)
1991	34	42	76	103	22	155 (40)	192 (47)	347 (53)
1992	-	-	-	-	-	-	-	-
1993	23	7	30	59	2	475 (630)	144 (197)	619 (287)
1994	29	21	50	36	3	273 (243)	198 (177)	471 (191)
1995	41	28	69	71	8	332 (140)	227 (99)	559 (155)

the summer. The number of recaptures varied by year and was generally low, leading to less precision and possible bias in some cases (Seber, 1982). The variation in proportion of females from year to year may be real or may be due to changes in the efficiency and experience of field observers in different seasons, when noting behaviours and calluses.

Between-season (open) population analysis

Goodness-of-fit and variance inflation

Attempts to apply the goodness-of-fit tests separately to the sub-groups were not successful due to sparse data. This sparsity arises because the recapture rate between samples is low, as is evident in the summary of the data shown in Table 2. The tests applied to the pooled data gave $\chi^2 = 28.23$, $df = 11$ and $P < 0.01$. The model therefore fails the goodness-of-fit test. There does not appear to be a systematic pattern of deviation from expected values in the tables.

As discussed by Lebreton *et al.* (1992) and Anderson *et al.* (1994), lack-of-fit may be due to structural deficiencies in the model, the presence of overdispersion in the data, or both. One cause of the lack-of-fit is the known heterogeneity between the sub-groups, which could not be tested separately for goodness-of-fit. A second is the social behaviour of sperm whales, which leads us to expect some overdispersion. Indeed, correlation between some individuals was evident in the observed distribution of capture histories.

Based on experience, Anderson *et al.* (1994) state that, 'Once one has found an adequate model structure, overdispersion c seems often to be just above one to as much as three'. The value of c in this study is:

$$\hat{c} = \frac{28.23}{11} = 2.57$$

which lies within the range suggested. Therefore, the assumption is made that this lack-of-fit is attributable to overdispersion, while emphasising that this is an assumption and structural problems with the model cannot be discounted.

Model selection

Fig. 4 schematically displays the path taken during the decision-making process from the most general JS model (A) to the most parsimonious model (G).

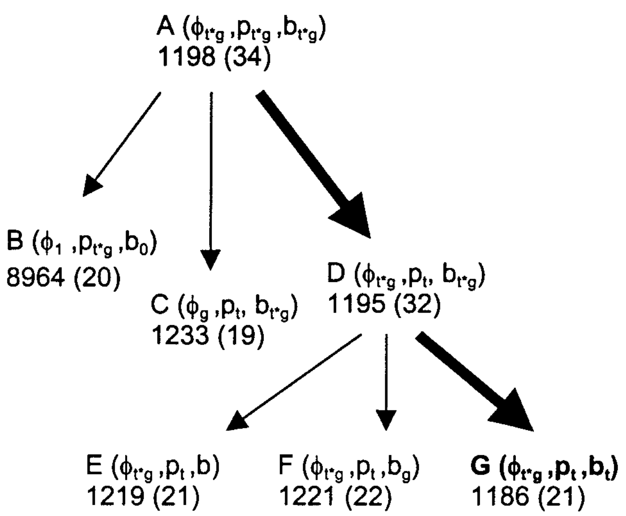

Fig. 4. Path of decisions made during model selection using QAIC. Letters index the model. The QAIC value is shown underneath each model with the number of parameters (total number parameters – number constraints) in parentheses. Heavy arrows show steps to preferred models, with model G finally selected.

Step A to B

A closed model fits very poorly compared to the full, open JS model. Select A.

Step A to C or D

Assume plausible model has time-varying capture probability (due to varying effort over the years) and survival which varies with group (see Table 2). Compare constant survival within groups versus time-varying survival within groups. Select D.

Step D to E, F or G

Constraints now placed on the birth rate (constant, time-varying or group varying). Select G, time-varying births.

The closed model B, in which there is no loss and no gain in the population, has a very poor relative fit compared with all other models examined (A, C-G). This is to be expected in a large population of mobile animals and as more members of the population are identified over the seasons. The final model is somewhat as expected. Survey effort varied from sample to sample (Table 1), so a time effect is included for capture probabilities. Survival in females is higher (Table 2), so a group effect is expected. The time-varying numbers of births and survival rates are in all likelihood due to immigration and emigration, as opposed to mortality or reproduction, which are both low in this long-lived animal (Lockyer, 1981; Best *et al.*, 1984).

Parameter estimates

Abundance estimates for the selected JS model are plotted in Fig. 5. No estimates are available for N_1 or N_K because the underlying parameters are non-identifiable. The standard errors are correlated with the estimates, a known property of JS estimates. Abundances appear to be more consistent within periods 1988-1990 and 1991-1994. The results suggest an increase in the population between 1990 and 1991 by a factor of about three.

Table 4 shows the estimates of survival and births for the two groups. As expected from Table 2, the survival rate for females is usually higher. The increase in abundance in 1991 (Table 4) corresponds to high numbers of births in the 1990 sample. The estimates are variable and sometimes required constraint to above zero (births) or below one (survival).

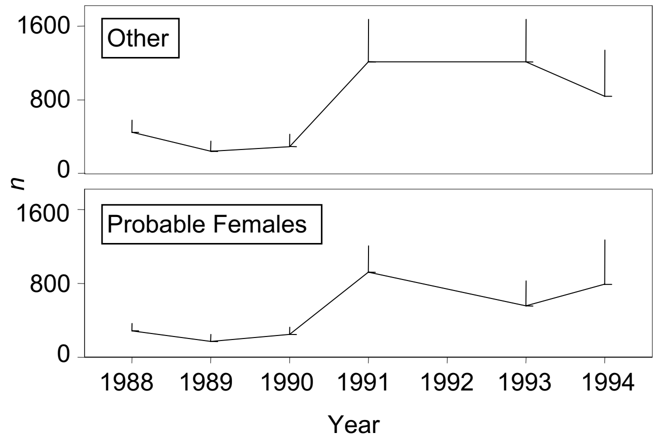

Fig. 5. Abundance estimates of the two sub-groups, with inflated upper standard error shown.

Table 4

Survival and birth estimates for selected JS model (ϕ_{t*g}, p_t, b_t). '1*' indicates that the estimate was constrained from above one, and '0*' that the estimate was constrained from below zero. No data collected in 1992.

	Probable females				Other			
Year	ϕ_F	se.$\sqrt{c}$	b_F	se.$\sqrt{c}$	ϕ_O	se.$\sqrt{c}$	b_O	se.$\sqrt{c}$
1988	0.60	0.22	0*	-	0.54	0.22	0*	-
1989	1*	-	75	64	0.71	0.34	119	102
1990	1*	-	675	274	0.46	0.30	1,078	429
1991	0.60	0.22	0*	-	1*	-	0*	-
1993	0.92	0.70	282	222	0.32	0.26	450	352

DISCUSSION

Validity of assumptions

Some major relevant assumptions of the models used are discussed below.

(1) *Random sampling*

In the present study, regular, relatively uniform sampling was approximated within a relatively small area around the islands of Pico, Faial and São Jorge, but was less methodical over the wider area around the Central Group of islands. Apart from this, the survey vessels did not concentrate effort in any particular area within or between seasons.

(2) *Population closure (Petersen estimates)*

The primary assumption made for these estimates is that the population is closed (both geographically and demographically). The photo-identification data in this study suggest that some whales are resighted over periods of months around the islands, that is, entire sampling periods, but the majority of whales were seen over at most a few days. It is not known whether the failure to resight is due to emigration or failed capture.

(3) *No re-immigration/temporary emigration (JS estimates)*

The closed population model applied over the full period (nine years) of this study fitted significantly less well than 'open' models, illustrating, as expected, that in the long term the assumption of closure is not valid.

It seems likely that the open population models in this study faced the same problem as some other cetacean mark-recapture studies (e.g. Calambokidis *et al.*, 1990; Hammond *et al.*, 1990), namely that temporary emigration occurs (violating an assumption of the JS model) because sampling is being carried out from only a small part of a roaming population. The study area is undoubtedly only a small part of the range of these whales, and does not even encompass the entire archipelago. The presence of temporary emigration means that from sample to sample marked individuals missing from the study area are not 'at risk' of capture. This is a form of heterogeneity of capture probabilities (a component of zero values is introduced in each sample). How this violation affects the estimates depends on the way in which the population visits the study area, and is further discussed below.

(4) *Homogeneity of capture*

Hammond (1990, p.135) writes that 'in practice, unequal catchability [is] likely to be a fact of life in all photo-identification studies of whales'. The degree of individual heterogeneity can be investigated in closed models with several samples (e.g. Hammond, 1990), but in this study the numbers of captures within seasons were generally not sufficiently high for this type of analysis. A measure taken here to reduce heterogeneity was stratification into two sub-groups with different catchability. The higher resighting rate of the 'probable females' sub-group may be due to greater site-fidelity, but could also be explained by a tendency for observers to preferentially photograph these animals, or a correlation with other attributes affecting their sightability. If for example, callused whales are more likely to be mature females than non-callused whales, they might be more approachable and less easily disturbed, and they are also more likely to have calves, which could constrain them to make shorter dives and thus fluke up more often. Also, callused females are likely to be older and therefore better-marked. It is important to bear in mind that assignments to sub-groups could not be made with any certainty (in contrast to some other mark-recapture studies, e.g. Darocher and Stirling, 1995).

(5) *Independence of captures*

Photographs taken from groups of sperm whales are clearly not independent, given the persistence of long-term associations between many individuals. This is assumed to decrease the precision of estimates, but not to bias them greatly. Estimates of variance were inflated here to reflect this problem but this technique is an area of ongoing statistical research. The approach of Whitehead *et al.* (1992; 1997) to the problem of non-independence in analysis of sperm whale photo-identification data was to calculate confidence limits by Monte Carlo simulation, based on assumed permanent units. This method may be preferable when stable groups are identifiable. However, it has not been applied here because sperm whale group structure is dynamic and identification of groups can be somewhat arbitrary (e.g. relying on cluster analysis) meaning that the assumption of fixed groups may not be a fair one.

Population definition

The open model estimates in this study refer to the population that visits the Azores, but current knowledge is such that it is not clear what that population is. Dufault *et al.* (1999) review suggestions that the North Atlantic population is a single stock but the information is poor and mainly based on information from three long range movements of tags and harpoons on male whales (Azores and Nova Scotia to Spain, Azores to Iceland). Females may have a more discrete population structure.

The migration habits and site-fidelity of the two main components of the population visiting the Azores (mature or maturing males and mixed groups) are not well understood. Although Clarke (1956) suggested that mixed groups might migrate to the Azores from breeding grounds to the south and spend the summer around the islands, there is little evidence for (or against) this suggestion. Some mature or maturing males may be found in temperate and tropical waters, including the Azores, in winter. However, recent findings seem to emphasise our ignorance of population movements. According to Avila de Melo and Martin (1985), catches of mature males peaked in the summer simultaneously with those of Iceland, so that

> 'a substantial part of this segment of the population remains throughout the year in warm temperate waters. Many, possibly most, mature males thus do not undertake an annual migration to and from higher latitudes'.

Based on strandings and sightings data from the North Atlantic, Evans (1997) concludes that since the mid-1970s 'an increased proportion of immature [males] [have been] migrating to high latitudes where some remain overwinter'.

Smith *et al.* (1999) describe a mark-recapture study of humpback whales carried out on the scale of the ocean basin. Such large-scale collaborative studies give a better understanding of migration patterns and the population structure, with an associated improvement in models and estimates. The recent establishment of the North Atlantic and Mediterranean Sperm Whale Catalogue (*http://www.ifaw.org/NAMSC*), which includes the images used in this analysis, is an encouraging step in this direction.

Population size

The Petersen estimates given here are relatively consistent from year to year, suggesting between 300-800 marked female and immature whales are found around the Central Group of the Azores in the summer. The open population models gave abundances of between about 400-800 between the years 1988-1990, increasing by a factor of three to between 1,600-2,200 animals between the years 1991-1994. The closed population estimates are of the animals in the study area in each season, while the open population estimates are of the more widely dispersed population, which visits the study area over several seasons. The fraction of unmarked whales is estimated at about 12%, so the abundance of both marked and unmarked whales may be estimated by increasing the Petersen and JS estimates by a factor of 1.14.

The study area forms only a part of the Azores archipelago and only a very small part of the range of the population, it is thus likely that there is temporary emigration from the study area. How temporary emigration affects the estimates is dependent on how the animals transit the study area. If there is random movement between the population in the study area and a population in some wider region, then the study may be regarded as representatively sampling the population in that wider region. The JS estimates are then approximately unbiased estimates of the wider population associated with the study area (Kendall *et al.*, 1997). On the other hand, if movements through the study area are in some way heterogeneous, then the estimates are biased in a way that cannot be assessed without further information.

Another consequence of the small study area relative to the population range may be that roaming individuals and groups come into the study area infrequently. The increase in abundance in 1991 is clearly not due to population growth, as sperm whales do not reproduce at such rates (Best *et al.*, 1984). The best interpretation of this increase may be that the composition of the population visiting the Azores changed at this time. There is a decrease in the proportion of recaptures in 1991 in the data. The results of the model suggest that a previously unmarked part of the population arrived in the region in 1991. The presence of new visitors does not seem to have caused any discernible increase in the summer density around the Azores, however.

Leaper *et al.* (1992) calculated density estimates with four surveys conducted around the central islands of Pico, São Jorge and Faial using acoustic survey methods in 1988 and 1989. Table 5 shows density estimates by the closed population mark-recapture and acoustic methods. The closed population estimates are used because the area of the population is better defined, whereas the area inhabited by the population estimated by the open population models is not. The area of sea within the study area is 28,378 n.miles^2. The estimates are quite similar, and although this is not a rigorous comparison (for example, the surveyed areas overlap but are not identical), the similarity in density estimates is encouraging.

Table 5

Density estimates by mark-recapture (closed population estimates) and by acoustic surveys.

Year	Density n.mile^{-2} (this study)	Density n.mile^{-2} (Leaper *et al.* 1992)
1988	0.0266	0.004, 0.013, 0.053
1989	0.0102	0.008, 0.01
1990	0.0189	
1991	0.0122	
1993	0.0218	
1994	0.0166	
1995	0.0197	

The Azores sperm whale catch varied between 400-741 per annum between the years 1937-1954 (Clarke, 1956). Most of these were caught in summer. The scale of this take could not have been sustained under the present day estimate of about 2,500 or less, indicating that the numbers visiting the Azores at that time may have been substantially larger. Movement patterns in the mid-20th century may also have differed, so whaling on the local population could have been replenished by the immigration of whales from the wider area.

ACKNOWLEDGEMENTS

The International Fund for Animal Welfare funds the work of *Song of the Whale*. We are grateful to the Ministry of Fisheries in the Azores for granting permits to work in Azorean waters, and to the Department of Oceanography and Fisheries of the University of the Azores for much help and cooperation over the years. The fieldwork would have been impossible without the hard work and dedication of a large number of sailors, scientists and student interns, too numerous to mention individually. However, we would particularly like to acknowledge Russell Leaper, Richard McLanaghan and Kit Rogers, who skippered *Song of the Whale* at various times. This paper is a much-enlarged and improved version of unpublished paper SC/51/CAWS27 presented to the IWC Scientific Committee, May 1999. Dr H. Whitehead (Dalhousie University) kindly provided comments and advice on the first submitted manuscript, and Russell Leaper on the second. We are also grateful to three anonymous reviewers for providing helpful remarks. One of these gave encouraging and extremely constructive comments, which greatly improved the analysis.

REFERENCES

Anderson, D.R., Burnham, K.P. and White, G.C. 1994. AIC model selection in overdispersed capture-recapture data. *Ecology* 75:1,780-93.

Arnason, A.N., Schwarz, C.J. and Boyer, G. 1998. *POPAN-5: A Data Maintenance and Analysis System for Mark-Recapture Data*. Department of Computer Science, University of Manitoba, Winnipeg, Canada. viii+265pp.

Arnbom, T. 1987. Individual identification of sperm whales. *Rep. int. Whal. Commn* 37:201-4.

Avila de Melo, A.M. and Martin, A.R. 1985. A study of male sperm whale length data from the Azorean and Madeiran catches, 1947-82. *Rep. int. Whal. Commn* 35:209-15.

Best, P.B., Canham, P.A.S. and MacLeod, N. 1984. Patterns of reproduction in sperm whales, *Physeter macrocephalus*. *Rep. int. Whal. Commn* (special issue) 6:51-79.

Calambokidis, J., Cubbage, J.C., Steiger, G.H., Balcomb, K.C. and Bloedel, P. 1990. Population estimates of humpback whales in the Gulf of the Farallones, California. *Rep. int. Whal. Commn* (special issue) 12:325-33.

Childerhouse, S.J. and Dawson, S.M. 1996. Stability of fluke marks used in individual photoidentification of male sperm whales at Kaikoura, New Zealand. *Mar. Mammal Sci.* 12(3):447-51.

Childerhouse, S.J., Dawson, S.M. and Slooten, E. 1995. Abundance and seasonal residence of sperm whales at Kaikoura, New Zealand. *Can. J. Zool.* 73(4):723-31.

Clarke, R. 1956. Sperm whales of the Azores. *Discovery Rep.* 28:237-98.

Clarke, R. 1981. Whales and dolphins of the Azores and their exploitation. *Rep. int. Whal. Commn* 31:607-15.

Darocher, A.E. and Stirling, I. 1995. Estimation of polar bear population size and survival in western Hudson Bay. *J. Wildl. Manage.* 59:215-21.

Dufault, S. and Whitehead, H. 1995. An assessment of changes with time in the marking patterns used for photoidentification of individual sperm whales, *Physeter macrocephalus*. *Mar. Mammal Sci.* 11(3):335-43.

Dufault, S., Whitehead, H. and Dillon, M. 1999. An examination of the current knowledge on the stock structure of sperm whales (*Physeter macrocephalus*) worldwide. *J. Cetacean Res. Manage.* 1(1):1-10.

Evans, P.G.H. 1997. Ecology of sperm whales (*Physeter macrocephalus*) in the eastern North Atlantic, with special reference to sightings and strandings records from the British Isles. *Biol. Unserer Zeit* 67-Suppl.:37-46.

Gordon, J.C.D. 1987. The behaviour and ecology of sperm whales off Sri Lanka. Ph.D. Thesis, University of Cambridge. 347pp.

Gordon, J. and Steiner, L. 1992. Ventilation and dive patterns in sperm whales, *Physeter macrocephalus*, in the Azores. *Rep. int. Whal. Commn* 42:561-5.

Hammond, P.S. 1986. Estimating the size of naturally marked whale populations using capture-recapture techniques. *Rep. int. Whal. Commn* (special issue) 8:253-82.

Hammond, P.S. 1990. Heterogeneity in the Gulf of Maine? Estimating humpback whale population size when capture probabilities are not equal. *Rep. int. Whal. Commn* (special issue) 12:135-9.

Hammond, P.S., Sears, R. and Bérubé, M. 1990. A note on problems in estimating the number of blue whales in the Gulf of St. Lawrence from photo-identification data. *Rep. int. Whal. Commn* (special issue) 12:141-2.

Jolly, G.M. 1965. Explicit estimates from capture-recapture data with both death and immigration - a stochastic model. *Biometrika* 52:225-47.

Katona, S.K. and Beard, J.A. 1990. Population size, migrations and feeding aggregations of the humpback whale (*Megaptera novaeangliae*) in the western North Atlantic Ocean. *Rep. int. Whal. Commn* (special issue) 12:295-305.

Kendall, W.L., Nichols, J.D. and Hines, J.E. 1997. Estimating temporary emigration using capture-recapture data with Pollock's robust design. *Ecology* 78:563-78.

Leaper, R., Chappell, O. and Gordon, J. 1992. The development of practical techniques for surveying sperm whale populations acoustically. *Rep. int. Whal. Commn* 42:549-60.

Lebreton, J.-D., Burnham, K.P., Clobert, J. and Anderson, D.R. 1992. Modeling survival and testing biological hypotheses using marked animals: a unified approach with case studies. *Ecol. Monogr.* 62(1):67-118.

Lockyer, C. 1981. Estimates of growth and energy budget for the sperm whale, *Physeter catodon*. *FAO Fish. Ser. (5) Mammals in the Seas* 3:489-504.

Martin, A.R. 1981. Further analysis of the post-war Icelandic sperm whale catch. *Rep. int. Whal. Commn* 31:765-8.

Martin, A.R. 1982. A link between the sperm whales occurring off Iceland and the Azores. *Mammalia* 46:259-60.

Pollock, K.H., Hines, J. and Nichols, J.D. 1985. Goodness-of-fit tests for open capture-recapture models. *Biometrics* 41:399-410.

Pollock, K.P., Nichols, J.D., Brownie, C. and Hines, J.E. 1990. Statistical inference for capture-recapture experiments. *Wildl. Monogr.* 107:1-97.

Schwarz, C.J. and Arnason, A.N. 1996. A general methodology for the analysis of open-model capture-recapture experiments. *Biometrics* 52:860-73.

Seber, G.A.F. 1982. *The Estimation of Animal Abundance and Related Parameters*. 2nd Edn. Charles Griffin and Company Ltd., London. i-xvii+654pp.

Smith, T.D., Allen, J., Clapham, P.J., Hammond, P.S., Katona, S., Larsen, F., Lien, J., Mattila, D., Palsbøll, P.J., Sigurjónsson, J., Stevick, P.T. and Øien, N. 1999. An ocean-basin-wide mark-recapture study of the North Atlantic humpback whale (*Megaptera novaeangliae*). *Mar. Mammal Sci.* 15(1):1-32.

Whitehead, H. 1990. Mark-recapture estimates with emigration and re-immigration. *Biometrics* 46:473-9.

Whitehead, H. and Gordon, J. 1986. Methods of obtaining data for assessing and modelling sperm whale populations which do not depend on catches. *Rep. int. Whal. Commn* (special issue) 8:149-65.

Whitehead, H., Waters, S. and Lyrholm, T. 1991. Social organization of female sperm whales and their offspring: constant companions and casual acquaintances. *Behav. Ecol. Sociobiol.* 29:385-9.

Whitehead, H., Waters, S. and Lyrholm, T. 1992. Population structure of female and immature sperm whales (*Physeter macrocephalus*) off the Galapagos islands. *Can. J. Fish. Aquat. Sci.* 49(1):78-84.

Whitehead, H., Christal, J. and Dufault, S. 1997. Past and distant whaling and the rapid decline of sperm whales off the Galápagos Islands. *Conserv. Biol.* 11(6):1,387-396.