Transit station or destination? Attendance patterns, movements, and abundance estimate of humpback whales off west South Africa from photographic and genotypic matching

J. Barendse[1], P.B. Best[1], M. Thornton[1], S.H. Elwen[1], H.C. Rosenbaum[2, 3], I. Carvalho[2, 3, 5], C. Pomilla [3], T.J.Q. Collins[2, 4], M. Meÿer[6], R. Leeney[7].

[1]Mammal Research Institute, University of Pretoria, c/o Iziko South African Museum, P.O. Box 61, Cape Town, 8000 South Africa.

[2]Ocean Giants Program, Wildlife Conservation Society, 2300 Southern Blvd., Bronx, NY, 12 10460-1099, USA.

[3]Sackler Institute for Comparative Genomics, American Museum of Natural History, Central Park West at 79th street, New York, NY 10024, USA.

[4]Environment Society of Oman, P.O. Box. 3955, P.C. 112, Ruwi, Sultanate of Oman.

[5]Faculdade de Ciências do Mar e Ambiente – Universidade do Algarve, Campus Gambelas, 8000-139,

Faro, Portugal.

[6]Oceans and Coasts, Department of Environmental Affairs, Private Bag X2, Roggebaai 8012, South Africa.

[7]School of Marine Science and Engineering, Portland Square, University of Plymouth, Drake Circus, Plymouth PL4 8AA, Devon, UK.

*Corresponding authors: jaco.barendse@gmail.com, pbest@iziko.org.za

ABSTRACT

Humpback whales found off west South Africa (WSA) are known to display an atypical migration that may include temporary residency and feeding during spring and summer. At a regional scale there is uncertainty about how these whales relate to the greater West African Breeding Stock B as a whole, with evidence both for and against its division into two sub-stocks. A database containing sighting information of humpback whales intercepted by boat in the WSA region from 1983 to 2008 was compiled. It included a total of 1,820 identification images of ventral tail flukes and lateral views of dorsal fins. After systematic within- and between-year matching of images of usable quality, it yielded 154 different individuals identified by tail flukes (TF), 230 by left dorsal fins (LDF), and 237 by right dorsal fins (RDF). Microsatellite (MS) matching of 216 skin biopsies yielded 156 individuals. By linking all possible sightings of the same individuals using all available identification features, the periodicity and seasonality of 281 individual whales were examined. Sixty whales were resighted on different days of which 44 were between different calendar years. The most resightings for one individual was 11 times, seen in six different years, while the longest interval between first and last sightings was about 18 years. A resighting rate of 15.6% of whales at intervals of a year or more indicates long-term fidelity to the region. Shorter intervals of 1 – 6 months between sequential sightings in the same year may suggest temporary residency. The TF image collection from WSA was compared to TF collections from four other regions, namely Gabon, Cabinda (Angola), Namibia and the Antarctic Humpback Whale Catalogue (AHWC). Three matches were detected were between WSA (in late spring or summer) and Gabon (in winter), confirming direct movement between these regions. The capture-recapture data of four different identification features (TF, RDF, LDF and MS) from six successive subsets of data from periods with the highest collection effort (2001 – 2007), were used to calculate the number of whales that utilise the region, using both closed and open-population models. Since dorsal fins have never been used to estimate abundance for humpback whales, the different identification features were evaluated for potential biases. This revealed 9 – 14% incidence of missed matches (false negatives) when using dorsal fins that will result in an overestimate, while variation in individual fluke-up behaviour may lower estimates due to heterogeneity of individual capture probability, by as much as 57-69%. Taking into consideration the small dataset and low number of recaptures, the most consistent and precise results were obtained from a fully time-dependent version of the Jolly-Seber open-population model, with annual survival fixed at 0.96, using the MS dataset. This suggests that the WSA feeding assemblage during the months of spring and summer of the study period numbered about 500. The relationship of these whales to those (perhaps strictly migratory) that may occur here in other seasons of the year, and their links to possible migratory routes and other feeding or breeding areas remain uncertain.

Key words: Abundance; Breeding Stock B; capture heterogeneity; capture-recapture; Chapman’s modified Petersen estimate; humpback whale; Megaptera novaeangliae; migration; photo identification; Program MARK; site fidelity; west South Africa.

INTRODUCTION

The west coast of South Africa should function as a near-shore migration corridor for humpback whales *Megaptera novaeangliae* based on its mid-latitude geographical position and occurrence of such behaviour along the east coast of South Africa (Findlay and Best 1996) and at similar locations elsewhere in the Southern Hemisphere (Bryden 1985; Dawbin 1966). However, in the vicinity of Saldanha Bay (at about 33^{O}S) historic and more contemporary observations have shown humpback whales to display seasonal residency from October to February (Olsen 1914; Best *et al.*1995; Findlay and Best 1995). Most recently, a shore-based survey there with near-complete seasonal coverage (Barendse *et al.* 2010) has shown that the high relative abundances recorded during these spring and summer months did not correspond to the timing of expected migration peaks, but rather to aggregations of whales feeding on krill *Euphausia lucens* and other crustacean prey.

Humpback whales found in the south-eastern Atlantic are designated to the International Whaling Commission's (IWC) Breeding Stock B (BSB) (IWC 1998) as included in the 'Comprehensive Assessment' of the IWC Scientific Committee (IWC 2010) for Southern Hemisphere populations. This region, particularly the west coast of Africa south of the equator, was characterised by extremely high catches from 1908 to 1914 and episodic catches thereafter (Best 1994). The whales from BSB are thought to migrate primarily to Antarctic Areas II (60^{O} W to 0^{O}) and III (0^{O} to 70^{O} E) for the austral summer, especially to the so-called 'nucleus feeding area' located between 10^{O} W and 10^{O} E (Figure 1a; IWC 2010). Based on mitochondrial and more recently, nuclear genetic evidence of population sub-structuring (Carvalho *et al.* 2010; Pomilla 2005; Pomilla and Rosenbaum 2006; Rosenbaum *et al.* 2009) BSB has been divided into two breeding sub-stocks, B1 and B2, with the Walvis Ridge or Angola/Benguela Front at about 18^{O} S proposed as a possible boundary (IWC 2010). However, the majority of sampling to date has been limited to only two widely separated localities: on the breeding ground off Gabon (Collins *et al.* 2008) which is thought to represent BSB1, and off the west South Africa (WSA) that presumably belongs to BSB2. Since no breeding behaviour has been observed (or is expected to take place) in WSA, the actual geographical location of the breeding ground for BSB2 remains unknown, and the proposed northern boundary at 18^{O}S would be inconsistent with the sea surface temperature regimes found for other humpback whale breeding grounds (Rasmussen *et al.* 2007). The detection of 10 whales biopsied off both Gabon and WSA (Carvalho *et al.* 2010) using microsatellite genotyping (Palsbøll *et al.* 1997) has raised questions about the BSB sub-division. Given that the whole coastal region between about 7 – 30 ^{O}S, comprising the territorial waters of Angola and Namibia (Figure 1a) is more or less unsampled it remains difficult to construct a more conclusive population structure model for the region.

The shore-based observations presented in Barendse *et al.* (2010) do not add to the current understanding of how these humpback whales relate to others in the region as derived from the genetic structure and microsatellite matches between Gabon and WSA (see above), nor provide information on whether the same individuals appear off Saldanha Bay during any of the same seasons in different years, or an accurate measure of how many whales utilise the area as a feeding ground. Individual photo identification (Katona and Whitehead 1981) may help to address these questions. Humpback whales are individually recognisable from two physical features that may be readily photographed: (1) their tail flukes, which includes the trailing edge, and the occurrence of natural marks, scarring, and pigmentation of their ventral surfaces (Katona and Whitehead 1981; Mizroch *et al.* 1990); and (2), the lateral view of their dorsal fins, that takes into account the shape of the fin, the prominence and distribution of knuckles on the caudal peduncle, and any scarring or pigmentation on the fin and/or flank (Kaufman *et al.* 1987). Although the use of dorsal fins and lateral body markings has yielded successful matches (Gill *et al.* 1995), the more distinctive flukes are favoured for use in regional photo identification catalogues. Such catalogues have been employed widely to identify migratory links (e.g. Stevick *et al.* 2004), examine regional movement patterns and population structure (e.g. Calambokidis *et al.* 2001), and calculate population sizes (e.g. Straley *et al.* 2009).

We present here results from the most comprehensive photo identification and genetic collection to date from the west South Africa region in order to examine within- and between-year attendance patterns, and investigate inter-regional movements between WSA, Namibia, Gabon, and Antarctic Areas II and III by comparing all available tail fluke collections from these areas. Furthermore, although not specifically collected for this purpose, the type of capture-recapture data obtained from the within-region photographic and genotypic matching may be suitable for the calculation of abundance estimates (Hammond 1986; Hammond *et al.* 1990). We try to estimate the number of humpback whales that may feed in the area during spring and summer using different approaches, including capture-recapture methods on selected sub-sets of data using different identification features (tail flukes, right and left dorsal fins, and microsatellites). Both closed and open population models are used, as is the norm in many published abundance estimates for large whales, including humpbacks (e.g. Calambokidis and Barlow 2004; Larsen and Hammond 2004; Straley *et al.* 2009). To our knowledge, this is the first time dorsal fins have been used to calculate abundance for this species, in addition to the more favoured flukes. The exposure of the latter is known to vary for individual whales which may affect individual capture probability (Perkins *et al.* 1984, 1985), while dorsal fins are always exposed and more easily photographed (Gill *et al.* 1995). Therefore, we examine potential sources of capture heterogeneity, sampling bias, and error that may result from the use of dorsal fins vs. tail flukes as photographic identification features, using double-marked animals (i.e. identified by more than one feature). The results are compared and discussed in terms of the estimation method or model applied, and identification feature used.

MATERIAL AND METHODS

Data collection and sighting database

The sighting database and photographic catalogue were compiled from a number of data sources (Table 1), but as a minimum requirement for inclusion had to be collected from within the Exclusive Economic Zone (EEZ) of South Africa, west of Cape Agulhas (20° E). These included data from humpback whales encountered incidentally during research work directed at other cetacean species, or during routine multi-disciplinary scientific cruises in the region, in the years 1983 - 2008 (Figure 1b). It further included all boat intercepts made during the work reported in Barendse *et al.* (2010) in the years 2001-2003, and those from another study dedicated to humpback whales at Cape Columbine in 1993, described by Best *et al.* (1995) (Figure 1c). Although it was attempted throughout to photograph the ventral side of the tail flukes and both left and right sides of the dorsal fin, and from 1993, to collect a biopsy from every whale encountered, any whale was included for which at least one of these were taken/collected and the date (day, month and year) and locality (latitude and longitude) of the sighting known. In most cases additional data (including group size, composition and behaviour, SST, depth, and duration of encounter) were also collected. Discrimination between individuals in the field (and association of specific images/biopsy attempts with individuals) was aided by onboard notes and sketches of body features, and by recording all photographic (film roll/data card numbers and frames) and biopsy sampling effort for each individual. This information was later used in the database to associate identification features with specific individuals seen during a sighting.

Prior to 2004 most images were recorded on high speed (ISO 400 and higher) black-and-white or colour negative, and colour positive film using motor-driven 35 mm single lens reflex (SLR) cameras with 100-300 mm manual focus zoom lenses; from January 2005 onwards these were replaced by digital autofocus SLR cameras. Digital images, and film frames scanned at 600 dots per inch (dpi) were cropped to maximise the coverage of the area of interest (i.e. tail flukes, or dorsal fin plus caudal peduncle), and imported into the sighting database in the JPG file format. Black and white negatives were scanned according to an unpublished protocol (L. Santos-Tieder *et al.* 2003, National Marine Mammal Lab, USA). Each image was individually assessed for photographic quality and orientation of the subject and a score based on a 5-point scale assigned to each of these categories (1 = not useable, 2 = poor, 3 = fair, 4 = good, and 5 = excellent). Every tail fluke (TF) image was further classified according to its ventral pigmentation pattern (or 'type') on a scale from 1 - 5, where 1 is all white (no central black bar between the left and right flukes) and 5 all black (see Rosenbaum *et al.* 1995). Flukes were rated for the part visible above water, i.e. whole, left fluke only, right fluke only, and trailing/leading edge. An additional classification type '0' was introduced for TF where it was impossible to assign types 1 - 5, either due to the unfavourable orientation or partial obscuring of the subject, or where the tail flukes were severely scarred or mutilated due to injury, such as killer whale (*Orcinus orca*) bites. Images were also assigned a score from 1 – 5 for individual distinctiveness of the subject although this rating was not used in any of the present analyses.

Skin biopsies were collected using the *Paxarms* rifle system (Krützen *et al.* 2002). All biopsy heads were sterilised by flaming after use. Samples were placed into individually labelled cryogenic tubes filled with a NaCl-saturated, 20% dimethylsulfoxide (DMSO) solution and placed on ice bricks in a cooler box. At the end of each day all skin samples were stored in a domestic freezer (-5 ^{O}C) until they could be transferred to a -15 ^{O}C freezer at the laboratory in Cape Town. Processing of samples was carried out at the Sackler Institute for Comparative Genomics (American Museum of Natural History).

Within-region matching

The matching described below was done separately for each identification feature. Thumbnail (100 dpi) or medium resolution (200/250 dpi) copies of the original pictures for all useable images (i.e. with photo and orientation quality ratings of poor and better) were viewed on 38 – 48 cm (15 -19 in) thin film transistor (TFT) computer screens. Original (large format) images were viewed for final decision making. Tail flukes were compared by pigmentation type to reduce the number of possible comparisons, first to all images of the same type, and then to all images from the preceding and following types (e.g. type 2 was compared to types 1, 2, and 3). Type 0 flukes were compared to all available images from all other types. In the case of dorsal fins, each image was compared with every other image. Within-year matching was carried out first, i.e. checking for matches of the same individuals on different days in the same year. Once completed, representative images of individual whales from each year were compared in chronological order to those of the subsequent year in the database and matches identified. The processes of within- and between-year matching were repeated by a second person. Where a match disagreed, it was reviewed and a consensus decision made to accept or reject it. Once all matching was completed, the best image(s) available per individual and identification feature were selected for representation in the overall catalogue, and a unique identification number assigned per identification feature.

The methodology for genotyping using 10 microsatellite loci is detailed in Carvalho *et al.* (2010). Each biopsy was associated to an individual sighting incident by its original biopsy number. In the case of a positive match between two skin biopsies, the laboratory code assigned to the earliest collected sample was retained as the identification number for that individual.

Periodicity and seasonality of resightings

Although matching was carried out for each feature independently a maximum of four identification features, *viz.* tail flukes (TF), right dorsal fins (RDF), left dorsal fins (LDF) and microsatellite (MS) could be collected for an individual whale at any given encounter. Wherever a common identification feature was identified between two or more different sightings, these could be linked. Thus, a full sighting history could be built based on all matches made through all available identification features between different encounters, even though these were not all collected at every sighting. It is important to note that failure to positively link one feature to another for the same individual could result in missed matches between different sightings. The problems of having multiple separate records for the same animal in a combined feature catalogue were highlighted by Gill *et al.* (1995), especially when dealing with large numbers of individuals. However, given the small total number of humpback whales identified, we believe the use of combined identification features was warranted in order to optimize the sample size for the purposes of examining trends in the growth of the catalogue and attendance patterns.

Within- and between (calendar) year occurrences of resighted individual whales were examined using combined identification features (genotype and photos of usable quality) for the entire database. The time interval between the dates of first and last sightings (excluding the first day) was calculated for all individual whales that were resighted on different days, both within and between years. For whales sighted on successive days, the time between sightings was assumed to be one day, i.e. rounded up to 24 h. Between-year time calculations took leap years into account. The number of days between sequential sighting events was also calculated for each individual whale. The seasonality of resightings for the entire sighting database was examined by sorting them by month, and separated on the basis of their overall resighting histories, i.e. seen only once, resighted within years only, and resighted between different calendar years. Note that the latter may have included some within-year sightings, but were not included in the 'within-year only' category.

Between-region photographic matching

The representative images of 154 individual humpback whales identified by TF that resulted from the WSA within-region matching (see above) were compared to TF collections from four other regions (see Figure 1a for localities):

Cabinda – Twenty-five individual whales of which identification pictures of TF (45 images in total) were taken during September 1998 off Cabinda, Angola, around oil production platforms some 50 nautical miles south of Congo River mouth (Best *et al.* 1999) were compared to the WSA, Namibia and Gabon catalogues;

Gabon – A total of 1,297 individuals represented by 9,776 images collected from 2001-2006 was compared to the WSA and Cabinda images. The database, area of collection, and matching procedures are fully described by Collins *et al.* (2008);

AHWC (feeding Areas II/III) - The Antarctic Humpback Whale Catalogue (AHWC) is a compilation of almost 5,000 photographs (TF, LDF and RDF) taken by miscellaneous contributors, both by scientists and non-scientists since 1987. The images originate from regions throughout the southern hemisphere, and the overall aim of the AHWC is to investigate movements of humpback whales between the Southern Ocean and lower latitude waters through an internationally collaborative project (Allen *et al.* 2008). It is currently maintained by the College of the Atlantic

(Maine, USA) and is publicly available on the web-based photo-sharing platform *Flickr*® (http://www.flickr.com/ahwc). The photostream can be viewed as a whole, or by sets, using the search tool to select any combination of tags or text, such as TF pigment type or locality of picture (for example, the tag 'T1 areaIII' would display all images of type 1 from Area III) (Judy Allen pers. comm.). The type 0 is not used in the AHWC. A total of 186 images representing 130 individuals, tagged as being from Areas II and III, were compared to the WSA images.

Namibia – There is presently no formal humpback whale catalogue for Namibia, but images have been collected at Walvis Bay, (23°00'S, 14°30'E) during research cruises directed at Heaviside's (*Cephalorhynchus heavisidii*) and bottlenose (*Tursiops truncatus*) dolphins, or by dolphin- and whale-watching operators in winter (June – August) and summer (January – March) of the years 2008, 2009 and 2010. Preliminary sorting and matching of these yielded 35 individuals (61 images).

Images of both whole and partial TF of all quality ratings except 'not useable' were considered. The AHWC does not catalogue non-useable images as individuals (Judy Allen pers. comm.). No matching was conducted between the Gabon catalogue and the images from AHWC and Namibia. Representative images of each individual in one database were systematically compared to those of the other, bracketed by fluke type (as described above for within-region matching) to avoid mismatches due to the variable assignment of TF types. All matches were checked and confirmed by a second person.

Abundance estimates

Catalogue size adjusted for annual survival

For each of the four identification features, a measure of the absolute minimum abundance was derived from the number of individual whales contained in the respective databases. This was done similarly to the method used by Straley *et al.* (2009) where the number of whales ($\tilde{N}_x$) alive in any given year (x) is calculated by adding the number of unknown (or 'new') individuals identified in that year ($\tilde{n}_x$), to the number estimated to have survived from the preceding year ($\tilde{N}_{x-1}$), the latter adjusted by an annual survival rate (ϕ) (Equation 1). The term $\tilde{N}_{x-1}$ is the sum of $\tilde{n}_{x-1}$ and $\tilde{N}_{x-2}$ (again adjusted with ϕ) and so forth. No variance can be calculated.

$$\tilde{N}_x = \tilde{n}_x + \phi(\tilde{N}_{x-1}) \qquad (1)$$

The value for ϕ was set at 0.96 as calculated for humpback whales in the North Pacific (Mizroch *et al.* 2004). Although this value is probably lower for non-adults (see Zerbini *et al.* 2010 for discussion), it is considered a reasonable estimate for annual adult survival, given that the area is not a breeding ground and very few calves were seen (Barendse *et al.* 2010).

Data selection for capture-recapture estimates

The only time period for which sufficient data were available for several years in sequence, and offered adequate seasonal coverage to permit estimation of abundance for whales that engage in spring/summer feeding, occurred during 2001 – 2007 (Table 2). This included the sighting data from the boat-based component of the work described in Barendse *et al.* (2010) (see above) and the remainder, of humpback whales encountered during work on feeding southern right whales (2003 – 2007) at Saldanha Bay (in September), and St Helena Bay (in October - December, rarely January) – note that this study had no shore-based watch (see Table 1). By restricting the data sub-sets to only certain seasons, the possible heterogeneity in capture probability introduced by different seasonal attendance patterns of individuals should be reduced. Six successive capture occasions (j) of six months each were identified, starting in September of one year and ending in February the following year (e.g. j_1 = 01 September 2001 to 28 February 2002, both dates inclusive) (see Appendix I).

Variation in photographic quality and the distinctiveness of natural marks can affect the ability to correctly match different photographs of the same individual, and hence the likelihood of a successful resighting (Friday *et al.* 2000; Gunnlaughsson and Sigurjónsson 1988; Hammond 1986; Stevick *et al.* 2001). For example, on images of poor quality, highly distinctive individuals may still be identified while matches of less distinctive animals are more likely to be missed (i.e. an increased probability of false negatives). To reduce such errors we applied the commonly used approach (e.g. Cerchio 1998; Straley *et al.* 2009; Friday *et al.* 2008) of excluding images below a certain quality; in this case those of quality and/or orientation rating of 'poor' and 'not useable' were not used for capture-recapture calculations and no partial TF pictures (halves or trailing edges) were included.

Closed population model

The two-sample Chapman's modified Petersen (CMP) estimator (Seber 1982) has been used elsewhere to calculate the size of feeding aggregations of humpback whales (e.g. Larsen and Hammond 2004; Straley *et al*. 2009). When applied over relatively short time periods (e.g. one-year intervals), it is considered an acceptable approach for a long-lived mammal with relatively low rates of natural mortality and recruitment, despite such populations generally not meeting the assumptions of closed population models. These assumptions (adapted from Seber 1982), applicable when using natural marks, are: (1) a constant population during the sampling period (no immigration/emigration, or births/deaths); (2) no loss of marks between sampling periods; (3) all marks are correctly recorded; (4) all whales have an equal chance of being recorded in the first sample; (5) both previously identified and newly sighted whales have equal probability of recapture in subsequent samples.

We employed the CMP estimator here due to its relative simplicity, and to illustrate issues that relate to different identification features used (see later), with the formula (Seber 1982):

$$N^* = \frac{(n_1+1)(n_2+1)}{(m_2+1)} - 1 \quad (2)$$

where N^* = estimated population size, n_1 = the number of whales identified during j_1, n_2 the number of whales identified during j_2, and m_2 the number of whales identified (i.e. matched) in both periods. The estimated variance (*v or* vâr) of N^* and the estimated coefficient of variation (CV*) of N^* were calculated according to formulas in Seber (1982):

$$\hat{\mathrm{var}}(N^*) = v = \frac{(n_1+1)(n_2+1)(n_1-m_2)(n_2-m_2)}{(m_2+1)^2(m_2+2)} \quad (3)$$

and

$$\mathrm{CV}^* = \sqrt{\hat{\mathrm{var}}(N^*)} \Big/ N^* \quad (4)$$

Confidence intervals (95%) for the CMP estimator were calculated with the log-normal transformed method as proposed by Burnham *et al.* (1987).

$$r = \exp\left(1.96\sqrt{\ln(1+(\mathrm{CV}(N^*))^2)}\right) \quad (5)$$

The lower confidence interval (CI) was calculated by dividing N^* by r, and the upper by the product of N^* and r.

The CMP calculation was restricted to the first pair of capture periods ($j_1 - j_2$) as these were the only ones with the primary effort directed at humpback whales, had the largest sample sizes, and where recaptures were detected for all identification features. Furthermore, sampling during j_1 - j_2 occurred at the same site of limited extent (i.e. within ±25 km radius from North Head, Saldanha Bay); this should reduce capture heterogeneity, a factor not accounted for by the CMP estimator between individuals, or over time (Hammond 1986). Such heterogeneity is regarded as highly likely to be a factor for all natural populations, resulting in underestimation of the true size of the population, sometimes considerably (Carothers 1973).

Open population models

Maximum-likelihood models of the Jolly-Seber (JS) type (Jolly 1965; Seber 1965; Schwarz and Seber 1999) are frequently used when the assumption of population closure is unlikely to be met, and when data from multiple capture periods are available. The POPAN option, included in the software Program MARK 5.1 (White and Burnham 1999; Schwarz and Arnason 2006) is one of the JS model formulations most readily available to biologists (Arnason and Schwartz 1999). It has therefore enjoyed wide application for generating population estimates from photographic and genotypic capture-recapture data for several cetacean species (e.g. North Pacific right whales *Eubalaena japonicus* - Wade *et al.* 2011; Indo-Pacific bottlenose dolphins *Tursiops aduncus* - Reisinger and Karczmarski 2010; killer whales - Reisinger *et al.* 2011), including humpback whales (Larsen and

Hammond 2004), and other large marine fauna such as whale sharks *Rhincodon typus* (Meekan *et al.* 2006); in some cases for very small populations for which limited data are available.

The POPAN model estimates the following parameters: the super-population size *N*; the apparent survival rate ϕ; the probability of entry into the population, or '*Pent*' with the alternative notations of *b* or β (the latter used here); and capture probability (*p*) at capture occasion *j* (Schwarz and Arnason 2006). The prescribed link functions (GC White, Program MARK Help files) namely, the Logit link for ϕ and *p*, and multinomial Logit (MLogit) link for β were used. Different variations of the model were applied to datasets for six successive capture occasions (j_1 - j_6) for all four identification features (TF, LDF, RDF, and MS) including all parameters fixed (.), full time-dependence (t) for ϕ, β and *p*, and with ϕ fixed at the biologically realistic value of 0.96 (see above). While the β parameter accounts for the contribution of births to the overall entry rate (Arnason and Schwarz 1999), and though there are published annual rates of increase (ROI) available for humpback whales (see Zerbini *et al.* 2010) it was not attempted to fix this at a specific value, given that our data are not likely to be (fully) representative of a discrete breeding population. Selection of the best models was done using Quasi-Akaike's Information Criterion (QAICc), adjusted for small sample sizes as implemented in MARK (Cooch and White 2006).

Biases in abundance estimates derived from different photographic identification features

Given that dorsal fins have never been used to calculate abundances for humpback whales, their reliability as a naturally marked feature for this purpose is untested. It is expected that the use of identification features with less information or that are less distinctive would be more difficult to match, which can result in misidentification (Hammond 1986), as is the case for other species where dorsal fins are used (Gowans and Whitehead 2001). Therefore, we examine the incidence and effect of missed matches, when using dorsal fins. Furthermore, we assess the possible impact of variation in individual fluking behaviour (on estimates) as it is a known idiosyncratic behavioural feature (see Perkins *et al.* 1984, 1985) and there was a sense during the data collection that it was more difficult to photograph the flukes of some individuals, a notion reinforced by fewer individuals identified by this feature compared to dorsal fins (see later). While we acknowledge that the use of genotypes is not completely free from error and may cause an upward bias in abundance estimates due to misidentification of microsatellites (see Lukacs and Burnham 2005; Wright *et al.* 2009), detailed consideration of this issue is beyond the scope of this paper, although we did compensate for it where applicable or possible in the analyses below.

Tests for false negative rates

Microsatellites were used as an independent (non-photographic) identification feature and all individuals ($n = 32$) that were identified by this feature and resighted on different days, were used as the sample. For each capture occasion (day) it was assessed whether a specific photographic feature of useable quality (>poor) was recorded;

then, whether or not a specific feature confirmed the matches made by microsatellite. The sample size per identification feature was the number of times both a MS match and a photograph of the feature in question were available (‘matching opportunities’). Failure to detect a photographic match constituted a false negative. As a simple test to quantify the positive bias caused by the detected false negative error rate (*e*), the pair-wise CMP estimator (see above, Equation 2) was calculated for the applicable dataset, using the false-negative correction developed by Stevick *et al.* (2001). The identification events (*s*) per sampling period (*j*) were taken as the sum of every time a whale was identified as an individual, excluding same-day resightings, therefore assuming that the boat crew recognised such individuals in different groups on the same day. Thus, to correct for the higher-than-actual total number of whales ‘identified’ due to missed matches within each sampling period, the numbers of individuals identified during j_1 and j_2 (n_1 and n_2) are calculated as

$$n'_j = \frac{n_j - e.s_j}{1-e} \tag{6}$$

The number of individuals matched between these samples (m_2) was increased by the error factors to correct for missed matches between j_1 and j_2 in the following manner

$$m'_2 = \frac{m_2}{(1-e)} \tag{7}$$

A comparison of the resultant population estimates with the uncorrected estimates provided estimates of the magnitude (%) of overestimation.

Variation in recording of tail flukes for resighted whales relative to other features

All whales resighted on different days (n = 60) were used as the sample, and the identification features collected during intercepts on these different days were compared. First, the number of times TF were recorded (of any photographic quality) during all intercepts of resighted whales was compared to that of other features. Second, the frequency with which TF were recorded in the case of multiple resightings was examined. Third, the duration of intercepts where TF were recorded was compared to those where no TF were recorded. Finally, the probability of recording TF or dorsal fins (left or right) for an individual whale was calculated by counting the number of intercepts during which the feature was recorded and expressing it as a fraction of the total number of times that the resighted whale was intercepted.

Use of double marks

Here we used TF as one type of mark, and LDF, RDF and MS respectively as alternative marks. For the two adjacent sampling periods (j_1 and j_2), the n_1 consisted of animals that were identified by both TF and the other mark in question, i.e. double-marked animals. The n_2 consisted of the total number of whales identified by either TF, or the alternative mark in the following sampling period; with recaptures (m_2) being those double-marked animals that were identified by whatever feature was used for n_2. This approach is intended to compare the relative capture probabilities of the two marks used: if they are equal, then recapture rates (and by inference, abundance estimates) should be similar whichever feature is used for the second sample. During the calculation using the CMP estimator (Equation 2), an error correction factor (*e*) was applied to dorsal fins and MS similar to that described above (i.e. n_2 was adjusted downward and m_2 adjusted upward, after Stevick *et al.* 2001), but n_1 was left unadjusted because the animals were already identified without error from their TF. The correction factors used for dorsal fins were those calculated from LDF and RDF false negative tests (see below). When MS was used as alternative identification feature it was adjusted by the mean allelic error rate of 0.065 calculated for the samples collected off WSA (Inês Carvalho, unpublished data).

RESULTS

Range and seasonality of collection effort

Due to the *ad hoc* and variable manner in which much of the photographic and genetic data were obtained, effort is loosely defined here as ‘collection days’, i.e. any day on which such data were collected. There were only 28 such days from 1983 to 2000, compared to 108 over the next eight years (Table 2). The greatest number (and days with boat availability) of collection days occurred between 2001 and 2006 during the two studies highlighted earlier (at Saldanha Bay and St Helena Bay) and made the greatest overall contribution in terms of number of images and individuals identified after matching was completed (Table 1). Other notable periods of data collection were during the earlier study at Cape Columbine (Best *et al.* 1995) and incidental humpback sightings made during a project on Heaviside’s dolphins (described in Elwen *et al.* 2009). Collection days, as a proportion of days where a boat was deployed, ranged from 12.8% (in 2000), to a high of 38.8% in 1993, and most years were at around 20 – 30 % (Table 2). Overall, at least one collection day was recorded during any given month, but effort was not evenly distributed across seasons. The autumn and winter months (March – August) had the poorest overall coverage with 10 or less collection days per month, while spring and summer months (September – February) were better sampled. Most collection days occurred in November (*n* = 30), and fewest in June (*n* = 1) (Table 2). The spatial extent of miscellaneous data collection along the west coast was fairly extensive (approx. 700 km between the northern- and southernmost sites; Figure 1b); however, the majority of data were collected within a fairly limited area of about 1 x 1 degree latitude/longitude grid square, no further than 25 km from the shore (Figure 1c), and included the major study sites mentioned above.

Within-region matching

Sighting database/catalogue

The WSA catalogue up to February 2008 included a total of 1,820 images, made up of 510 TF, 694 RDF, and 616 LDF (Table 1), representing 446 individual sighting histories collected during 225 boat intercepts/encounters. Excluding images that were deemed ‘not useable’, 154 individuals were identified using only TF, 237 by RDF, and 230 by LDF (see Table 4). Microsatellite genotyping of 216 skin biopsies yielded 56 samples matched to one or more other samples, representing 156 individuals, three of which were identified by microsatellite only (i.e. were not photographed). By linking different individual identification features to common sightings, a total of 289 individual whales were identified with ‘combined features’, although eight only had ‘not useable’ images and were thus excluded (*n* = 281). Few animals (<10 per annum) were identified before the advent of dedicated field work in 2001 (Figure 2), when most individuals were identified in a single year (80). New additions remained at fairly high levels for the following five years (> 25 individuals per annum) although there was a steady decrease in the growth rate of the database (Figure 2).

Resighting rates, intervals and seasonality

Using combined identification features (*n* = 281), 214 individual whales were seen once only, seven were resighted on the same day (i.e. in more than one group), and 60 (21.35%) on different days. Forty-four whales were resighted between calendar years, the majority once only (30), followed by twice (7) to a maximum of five resightings (i.e. in six different years). Only 12 of these between-year sightings were not seen on multiple occasions in the same year, with one individual recorded a total of 11 times (the same whale that was seen in six different years).

The shortest interval between first and last sighting events was one day and the longest 18 years, with the mean interval being 3.4 yr and the median 1.5 yr. Most whales were resighted within one year (23), followed by a 1 - 2 yr interval (17). For 14 whales, the interval was longer than four years, and for six of these, longer than 12 years (Figure 3). A breakdown of time intervals between sequential sightings (Figure 4) of all resighted whales showed that most individuals were resighted on the same day (35 times), or within a week of the previous sighting. Resightings at intervals of more than a week, but less than six months, were relatively few (<10). The next most commonly observed resighting intervals were at 6 – 12 mo and 1 – 2 yr (Figure 4). Intervals of between 2 and 3 yr and longer than 5 yr were recorded less than 10 times each, while between 3 and 5 yr was not very common.

None of the 32 individual whales seen during winter months (June to August) were resighted (Figure 5). During all other months some of the whales seen were resighted on other occasions, the majority between calendar years. Between October and January, a small proportion of resighted individuals were same-year resightings only;

however, from February to May all resighted individuals were between years and 50% or more of whales seen during these three months had been seen previously (Figure 5).

Between-region matching

None of the images from Cabinda or Namibia matched a whale in any of the catalogues they were compared with. Three matches were made between the WSA and Gabon catalogues, and two between WSA catalogue and the Area II/III images contained in the AHWC (Table 3; also see Figure 1b). Three of these whales (TF-ZAW-01-005, TF-ZAW-03-017 and TF-ZAW-04-005) were also resighted in different years off WSA (Table 3). The matches with the AHWC were found to be with two humpbacks sighted together on the first day of the IWC-SOWER (Southern Ocean Whale and Ecosystem Research) cruise that departed on 22 December 2005 from Cape Town for the Antarctic; the images were inaccurately tagged in the database as being from Area III. Both were males (determined from biopsies collected off WSA) and the one animal (TF-ZAW-05-007) was seen less than a month before in St Helena Bay, some 150 km to the north (Table 3). The second animal had been seen previously in St Helena Bay in December 2004 when it was accompanying a cow-calf pair, and was identified as a possible yearling calf. It was also seen subsequently, on 22 November 2006 (also in St Helena Bay) with a different female, when several defecations were observed, presumably an indication of recent feeding.

A northward transit with duration of about 230 d between sequential sightings was recorded for two of the WSA-Gabon matches (Table 3). Shorter southward transit periods (40 - 80 d) between Gabon and WSA for sightings in the same calendar years were recorded. The one male (TF-ZAW-01-005) provides an interesting perspective in that it was sighted off WSA during the years before and after being photographed in Gabonese waters. It was first seen off Saldanha on 16 December 2001 as part of a group that defecated. On 6 August 2002 (233 d later) it was identified off Gabon, before appearing off Saldanha Bay 88 d later, where it apparently remained in the vicinity for a period of over two months, to be resighted on 14 January 2003. It was again resighted on 7 November 2006 when it approached the research boat during a plankton haul in St Helena Bay (Table 3).

Abundance estimates

Quality control criteria excluded 122 out of 1,409 images (all photographic features) from the datasets used for abundance estimates. The summary capture-recapture tables of data used in the models (given in Appendix I) show the overall small sample sizes and few recaptures, especially for TF. During the first two sampling periods more individuals were identified by dorsal fins than other methods, and more matches were made, although for the remaining four periods most whales were identified by MS. The latter feature generally had the highest recapture rate relative to total number of whales identified, and between all pairs of recapture periods. Fewer resightings were recorded for LDF than RDF.

Adjusted catalogue size

The total number of individual humpback whales in the database (after correction for annual survival) represented by TF was over 30% lower than for LDF or RDF, while almost the same as MS, bearing in mind that biopsy sampling only started in 1993 (Table 4). Given that failure to match dorsal fins (or genotypes) that belong to the same individual contained in the database would inflate the catalogue size for the relevant feature, the total catalogue size was reduced (after correction of survival) by the calculated false negative rate for the respective feature (see Table 7). The numbers of whales identified by LDF and RDF were still greater than for TF (by 22.4% and 29.7% respectively).

Closed population models

The CMP estimates from TF data were the lowest overall, even less than the lower 95% confidence intervals of the estimates for all other identification features (Table 5) and less than the adjusted TF catalogue size. Tail fluke estimates were between 70 and 80% lower than uncorrected dorsal fin estimates, and about half the uncorrected MS estimates. Even when these features were corrected for false negative errors (see later) the TF estimates were still 45 - 75 % smaller. The highest overall N^* was from RDF, then LDF and MS, although all estimates had fairly wide 95% CI's. The estimates from the genotypic recaptures had the lowest CVs.

Open population models

Model configurations with all or most parameters constant or fixed $\{\phi_{.}\ \beta_{.}\ p_{.}\}$ and $\{\phi_{0.96}\ \beta_{.}\ p_{.}\}$, or with capture probability set to vary between capture periods, and other parameters constant or fixed $\{\phi_{0.96}\ \beta_{.}\ p_{t}\}$ showed very poor fit, or failed to converge, and were not considered. The remaining model variants applied were:

(1) full time-variance for all parameters $\{\phi_{t}\ \beta_{t}\ p_{t}\}$;

(2) full time-variance for two parameters with ϕ fixed at 0.96 $\{\phi_{0.96}\ \beta_{t}\ p_{t}\}$;

(3) ϕ fixed at 0.96, β set to vary fully over time, and p constant $\{\phi_{0.96}\ \beta_{t}\ p_{.}\}$.

Using the $\Delta QAIC_c$ as indication, model 2 showed the best fit for all identification features except for the LDF data for which the full time-variant (model 1) fitted best (Table 6). For TF, model 3 with p fixed had almost equal support to model 2 ($\Delta QAIC_c < 2$) although yielding a considerably lower estimate for N. Model 3 was also fairly well supported for MS data although model 2 performed better (Table 6); estimates based on this feature showed the least variation between model variants, and the tightest confidence intervals. All other identification features showed considerable variation and very wide CI's, especially those derived from dorsal fins. For all identification features model 3, the variant with fewest parameters, yielded the lowest CVs, although it was not well supported (zero likelihood) for the dorsal fin data (Table 6). The estimates for TF (with the exception of model 2) were lower than for other features. The estimates derived from dorsal fins were about double the highest TF or MS estimates, but with high variance and CI's (Table 6). Goodness-of-fit tests available in MARK did not yield results due to inadequate data availability. Given the sparse data and low number of recaptures, no attempt was made to model more complex configurations.

Potential biases in abundance estimates for different photographic identification features

False negatives

Assuming that the microsatellite identifications were correct, photographs of LDF and RDF when used alone as an identification feature resulted in 13.8% and 9.1% missed matches respectively, whereas no missed matches were detected for tail flukes (Table 7). No false positives were detected for dorsal fins. To test for misidentifications using microsatellites, individuals resighted by tail flukes on different days using pictures of quality and/or orientation > 'poor' were used as a control (11 individuals, intercepted 24 times), and were compared to matches obtained by microsatellite (where biopsies were taken). No false negatives were detected in seven matching opportunities. The values for N^* for the LDF and RDF recapture data and corrected for by the respective error rates (0.138 and 0.091) were 18% and 10% lower than the respective uncorrected values (Table 5). Although no microsatellite mismatches were detected, an abundance estimate corrected for the mean allelic error rate (0.065) is included for comparison: it was 17% lower than the uncorrected estimate (Table 5).

Individual variation in fluke exposure relative to other features

For 21.67% of the whales resighted on different days (n = 60), no pictures of TF were collected, for 20% no biopsies, 3.33% no RDF, and 1.67% no LDF photographs. In the majority of cases, TF photographs (for the 47 whales) were obtained during the first intercept/encounter (65.96%), 27.66% during the second, and 6.38% during the third and fourth. Furthermore, during all intercepts involving these resighted whales (n = 183 - some whales were in the same groups), TF pictures were collected during only 57.4% of intercepts, compared to 92.9% for dorsal fins. There was no significant difference between the mean duration of intercepts where TF were photographed (73.84 min ± 3.88 SE, n=146) and where it was not (83 min ± 11.14 SE, n = 31) (t = -0.93, df = 175, p = 0.35). The probability of recording a dorsal fin image (right or left) of an individual whale, every time it was encountered was high (Figure 6). This was not the case for TF, where for individual whales the probability of recording this feature during all, half, or none of encounters was very similar (28, 25 and 23 % respectively) (Figure 6).

Double-mark models

For the selected capture periods used for CMP estimates, the models where TF was used for recapture (m_2) and the second sample (n_2) invariably resulted in lower abundance estimates than when the alternative features were used (Table 8). The highest estimates were calculated with RDF as the alternative mark, while LDF and MS yielded very similar estimates. Those in which TF were used for n_2 were 0.31, 0.39 and 0.43 of those using RDF, LDF and MS respectively for n_2.

DISCUSSION

Sighting database, resightings and migratory links

The distribution over space and time of contributions to the sighting database and resulting resighting rates reflected the generally low and often inconsistent collection effort. This makes it difficult to interpret the observed resighting rates relative to other capture-recapture studies with greater geographic coverage and higher sampling effort, such as is obtainable during dedicated survey cruises (e.g. Larsen and Hammond 2004; Wedekin *et al.* 2010) or simultaneous surveys from multiple platforms or sites (e.g. Smith *et al.* 1999; Calambokidis and Barlow 2004). Resighting rates are known to be much higher at feeding sites (some >50%) compared to breeding grounds (10% or less) (Herman *et al.* 2010). Our between-year resighting rate of 15.65% seems relatively high given the low effort, limited extent of sampling, and that it is neither a breeding nor typical feeding area, and could be indicative of strong site fidelity or a small ‘population’, or both. There is strong historic evidence for severe depletion of this assemblage of humpback whales: during the five years immediately prior to protection of the species in 1963, annual catches from the Donkergat whaling station at Saldanha Bay averaged only five whales taken during a six-month season (IWS 1964), compared to catches of 208-244 humpback whales a season over the first three years of whaling from Donkergat (Best 1994).

Long-term site fidelity is supported by the majority of individual resightings occurring at annual or biennial intervals, and on six occasions, up to a decade apart. While most of these whales were seen only twice, others were seen in three or more different calendar years up to a maximum of six different years; again, the *ad hoc* collection effort probably confounded the ability to detect more returns of known animals. Returns over such time scales may confirm fidelity to the area but do not necessarily identify it as anything other than a migratory corridor. However, this is challenged by sequential resightings of the same individuals on the same day, or within a week of the first sighting, suggesting that they were not merely moving through the area as expected during a typical migration. This confirms the phenomenon of temporary residency first observed during the 1993 study at Cape Columbine (Best *et al.* 1995) when the same 10 humpback whales were seen on average 2.4 times during a month-and-a-half long period. It is thus not inconceivable that whales resighted in the same year or breeding season at periods of 1 – 6 months apart could be moving around locally and remain in the general area for days, weeks, or even months. Continuous occupancy cannot be proven however, and it should be equally feasible for animals to depart from, and return to the area in such time-spans. Here, the matches with the two other regional catalogues shed more light on the possible nature of such movements. The resightings off Cape Town detected through the AHWC matches confirm the occurrence of local movements beyond the core study area, and a net southward movement during summer. The three matches between Gabon and WSA independently confirm evidence from genotypic matches (Carvalho *et al.* 2010) that showed exchange of individual whales between these two proposed ‘sub-stocks’. All three records indicate movement between winter (August, September) in Gabon and late spring/summer (November – February) in WSA, with the movement occurring in both directions but with the fastest transits from north to south. The monthly distribution of animals that were seen on more than one occasion (both within- and between-region resightings) suggests that humpback whales that engage in feeding during late spring, and in particular, summer months of the west coast of South Africa (as discussed in Barendse *et al.* 2010) are also likely to be encountered repeatedly during these months in other years. Furthermore, some of these whales were present off Gabon during August/September, so presumably overwintered there. On the other hand, the paucity of resightings of any of the animals identified off WSA during June to August (mid-winter) during which 11.38% of all whales identified were seen (based on combined features) suggests that whales present in the region at that time might belong to a different component of the population. If, as suggested by other evidence (see Chittleborough 1965; Dawbin 1966) this corresponds to the timing of an expected northern migration, it would appear that animals utilising the study area as a spring/summer feeding ground on their southward migration do not necessarily take the same route moving north to the breeding grounds.

The lack of any matches with the few available Namibian animals does not preclude the occurrence of a ‘typical’ coastal migration from headland to headland (Chittleborough 1965; Dawbin 1966) from WSA through Namibia to a more northerly destination, and back, on the southern migration. However, historical evidence does not support this in that off Namibia catches showed a sign of ‘recovery’ in 1925-30 after initial depletion while off WSA they remained very low (Best and Allison 2010). Possibly the coastal migration stream is cumulative rather than unitary, with northward-moving animals progressively converging on the coast with decreasing latitude, and southward-moving animals leaving the coast with increasing latitude. At Walvis Bay there were some sightings during late summer, but most humpback whale sightings were in winter (June, July, August) (S. Elwen, unpublished data); however, there was no research effort after these months to allow detection a later peak, as observed off WSA, although whale and dolphin watching operators did encounter some humpback whales during September and October. Furthermore, historical catches (Best and Allison 2010) showed strong bimodality at Walvis Bay, with a peak in June/July and another in October/November.

Abundance estimates

The available capture-recapture data were very limited in terms of sample sizes and number of recaptures detected between sampling periods: especially for TF that was apparently under-represented in the database at only 121 individuals (after correction for survival), while the similarly lower number of whales identified by MS could be partly attributed to its implementation (as means of identification) a decade later than photo identification. The higher numbers of whales identified by dorsal fins may reflect that they are more easily photographed: unlike TF, they are always exposed during surfacings (but see later), while biopsy sampling requires the closest approach of all sampling methods and may be difficult to achieve for boat-shy individuals that otherwise can be photographed. The CMP estimates for N^* during the first sampling periods more or less reflect this relative representation of features in the sighting database, i.e. dorsal fins the highest and TF the lowest at an improbable 67 individuals. For the remaining features, the estimates for N^* ranged between 122 (MS) and 265 (RDF) after error correction. That the TF estimate is an order of magnitude lower than for any other feature and is comparable to the adjusted catalogue size for TF at this time (69), suggests that the ability to capture a whale using TF may be affected by an additional factor (see later).

The poor (or non-) performance of the POPAN open-population models when capture probability (p) and probability of entry (β) were fixed, again suggests considerable variation in these parameters between capture periods; perhaps not surprising, given the low effort, the differences in sampling strategy (with or without a land watch) and location of sampled areas (Saldanha vs. St Helena Bay) between j_1 - j_2 and j_3 - j_6. For the most successful model variants apparent survival (ϕ) was fixed at 0.96 and both β and p were fully time-dependent (model 2), or β time-dependent and p fixed (model 3). Before looking at the magnitude of the estimated values for N (the super-population) by the POPAN models, it is worth considering it in the context of what was sampled. This parameter provides an estimate for the total number of animals, both captured or not, available in a (hypothetical?) super-population (GC White, Program MARK Help files). For example, in a study on bottlenose dolphins, Reisinger and Karczmarski (2010) using POPAN concluded that the N there potentially represented the dolphin population along a considerable segment, or even the entire South African coastline, while in other studies it has been taken to represent a full population of right whales (Wade *et al.* 2011), a sub-population of whale sharks (Meekan *et al.* 2006) or feeding assemblage of humpback whales (Larsen and Hammond 2004). From this it is apparent that N as estimated by POPAN is likely to represent more than simply the size of the feeding assemblage at Saldanha and St Helena Bays during any given season, but probably also includes whales that migrated through the area, or whose utilisation of the area varied between years and should probably be treated as an upper sub-population limit.

The abundances obtained from the open population models were larger than those from other methods. Similar to the other assessment methods, the dorsal fin estimates were twice as high as those for other features, most likely as a result of the occurrence of false negative errors (see later). Although a correction parameter for photographic and genotypic mismatches is not available in the POPAN model, estimates that include such errors would most likely still fall within the wide confidence intervals obtained, especially for dorsal fins. Tail fluke estimates were lower than for other features (but not as dramatically as with the CMP model) with the exception of model 2 where the estimate was very similar to that for MS, albeit with a much higher CV. The MS capture-recapture data yielded the most consistent estimates with the lowest variation between models, and tightest confidence intervals for individual variants. The model-averaged estimate of *N* for this feature (calculated from all three variants in MARK) was 510±143 SE (230 – 790, 95% CI's).

Biases as a result of different identification features

There are numerous known issues associated with the use of natural markings for abundance estimates (see Hammond 1986 for review), but (physical) loss of marks is not considered a major problem with humpback whales. Although dorsal fins are commonly used by researchers in the field to distinguish between different whales in a group while collecting data during boat encounters, and have been proposed as a potentially even more stable identification feature than ventral tail fluke pigmentation (Blackmer *et al.* 2000), this is, to our knowledge, the first time they have been used for an abundance estimate for the species. While false positives are probably rare in photo identification studies (we detected none), false negatives are thought to be more common, with a higher error rate the poorer the quality of pictures (Stevick *et al.* 2001). This is more likely to apply to dorsal fins, as they are smaller and have fewer distinguishing features than tail flukes. Although dorsal fin photographs (of sufficient quality) were not collected during all intercepts, there were no resighted individuals for which dorsal fin pictures were unavailable. There may be differences in the ability of researchers to obtain good quality images of these different identification features: during a typical approach from the rear, chances are good of obtaining a TF picture (provided that they are adequately exposed, see below). For dorsal fins, a considerable amount of manoeuvring of the boat is required to position the photographer at a right angle to the whale, while still at the surface. The angle between the camera and the whale affects the quality of dorsal fin pictures to a greater extent than for TF (J. Barendse pers. obs.) and poor photo quality can be the source of substantial heterogeneity in capture probability when using dorsal fins in other species e.g. northern bottlenose whales *Hyperoodon ampullatus* (Gowans and Whitehead 2001). The application of some quality control (e.g. removal of photos with incomplete subjects and those of low quality) has been shown to adequately reduce error rates when using TF (Friday *et al.* 2008), but it is unknown to what degree this is applicable to dorsal fins, and we were unable to test this with our small sample sizes.

Assuming the microsatellite identifications were correct, dorsal fin photographs when used alone as an identification feature resulted in 9-14% missed matches, whereas this does not appear to apply to the same extent for TF or MS. If left uncorrected, this may result in a substantial over-estimation (up to 18%) of abundance when using closed population models, and produce high estimates of low confidence in open population models. This conclusion however may be case-specific, depending to a large extent on data collection protocol, photographic quality, laboratory procedures, and the size of the catalogue. The differences between abundance estimates for RDF and LDF (although less pronounced compared to TF) suggests that there may have also been a difference in the ability of photographers to obtain useable images from both sides of an individual. The reason for this is not immediately apparent, although individual behaviour could contribute to such a bias. Clapham *et al.* (1995) reported strongly lateralised behaviour by humpback whales that apparently favoured their right-side during feeding and flippering behaviour; it is possible that whales could preferentially present their right side to the boat. However, we are unable to test this with the available data.

Relative to other identification features (even after they had been corrected for missed matches) TF yielded the overall lowest abundance estimates. This suggested that fluking as an individual behavioural trait could affect the probability of an individual being sampled, and was supported by the finding that for resighted whales, the probability of collecting TF pictures during all, half, or none of the intercepts was nearly equal. The fact that for all resighted whales, over 20% had no TF image collected at all during intercepts of similar mean duration, and that in the majority of cases (65%) flukes were photographed during the first intercept, suggests that fluking is an idiosyncratic feature for humpback whales in this area. The frequency of exposure of the ventral surfaces of the flukes is a behaviour known to vary with sex (Rice *et al.* 1987), age, reproductive and behavioural class, and group size, with fluking rates ranging from <10 per 100 dives for cows, calves and principal escorts, to 46.5 per 100 dives for single animals (Smith *et al.* 1999). If some whales consistently fluke less often than others, or do not fluke at all, the resulting heterogeneity of capture probabilities will lead to underestimation of population size (Barlow *et al.* 2011), such as is strongly suggested by these data. In West Greenland such (downward) bias was estimated at 10 – 20 % of the population size (Perkins *et al.* 1984), but presumably can vary with area, season, or photographic protocol. Based on the CMP abundance estimates for the double- marked whales, those using TF recaptures and identifications during the second sampling period were 57 - 69% lower than those when using an alternative feature. While this conclusion about the effects of individual fluking behaviour on population estimation may only be valid for the whales observed in some areas (such as WSA) as humpback whale behaviour may differ (and sampling protocol vary) in different parts of its range, the potential effects shown here are certainly large enough to warrant similar investigations in other areas. It is however more difficult to quantify (and thus correct for) heterogeneity attributable to individual behaviour compared to other sources of error (e.g. photographic quality) (Barlow *et al.* 2011).

Genotypic abundance estimates may be considered independent from those obtained from photo identification (as suggested by Gubili *et al.* 2009) with a lesser degree of bias from sampling heterogeneity applicable to photographs of natural marks (i.e. image quality and fluking behaviour). On a broader scale, because genetic and photographic sampling took place simultaneously and from the same platform, both could be considered subject to the same potential biases caused by non-representative sampling effort. The abundance estimates derived from MS recaptures were lower than those from dorsal fins, but more similar (but higher) to those obtained for TF, though more recaptures were made with MS. Heterogeneity of obtaining a biopsy (i.e. capture probability) cannot be excluded, given that it requires a closer approach than obtaining a TF image, for example, and recognising that for 20% of resighted whales no biopsy was collected. This could be as a result of a different individual behavioural response to boat approaches, as has been tested for whale watching boats (e.g. Scheidat *et al.* 2004; Stamation *et al.* 2010): the whales could exhibit boat avoidance (i.e. 'trap shy' animals) or be boat friendly ('trap-happy'), both of which will cause capture heterogeneity and bias in abundance estimates. However, there are certain issues relating to methodology and laboratory procedures that are specific to the use of molecular tags and may bias abundance estimates downward (Mills *et al.* 2000; Waits and Leberg 2000) or upward (Lukacs and Burnham 2005; Wright *et al.* 2009).

Conclusion

The consolidated photographic and genotypic sighting database for humpback whale from WSA contributes to a better understanding of residency rates and long-term attendance patterns to the region. It reveals that some whales routinely visit the coastal waters of the St Helena/Saldanha Bay region, showing high fidelity to a geographically limited area to participate in feeding during spring and summer. While acknowledging the effects of the sampling approach, low effort and small sample sizes on capture probability, possible structuring of the population would further contribute heterogeneity to individual capture, and the notion that there may be a strictly migratory (or 'transient') component is supported by the high number of once-off sightings, especially during winter. The situation is probably fairly complex, but similar to that observed off California where whales that share a wintering region show strong fidelity to specific feeding areas, with limited exchange between these, although whales from different sub-areas may make use of the same migratory corridors (Calambokidis *et al.* 2001). Given the limitations of the data, and violation of closure, the estimates from the closed-population model are likely to be underestimations of the total number of humpback whale found here on a seasonal basis, although their use did allow us to explore issues related to the use of different identification features. The findings that using dorsal fins and TF can cause substantial over-estimation and under-estimation respectively are important, especially when dealing with such a small dataset, and given that the ventral TF pattern is the standard identification feature used for humpback whales. Whether humpback whale dorsal fins are distinctive enough features to use in large catalogues (or as alternative identification feature) is debatable and it may be more appropriate to only include animals that are considered adequately 'marked', similar to the practise in dolphin studies (e.g. Elwen *et al.* 2009), although this approach may violate assumptions of equal capture probability, depending on the model used (see Reisinger *at al.* 2011). The open-population models with few restrictions and using multiple capture periods fitted the capture-recapture data better, although there was considerable variation between different identification features and model variants used. The most consistent estimates (and those with the narrowest confidence intervals) were obtained from the microsatellite data, putting the number of humpback whales that visited the area during the study period at about 500 animals, a value that falls within the confidence intervals for dorsal fins (lower range), and tail flukes (mid to upper range). The geographic extent of this estimate is not clear, nor whether all these animals feed in the region of Saldanha every season, or how they relate to whales along the rest of the West African coast. Although this assemblage does not exist in isolation from the greater Breeding Stock B (given the photographic matches), it does seem to represent a previously undescribed situation in the Southern Hemisphere, although the possibility cannot be excluded that such behaviour, or similar assemblages may occur at other localities where comparable oceanographic conditions (i.e. wind-driven upwelling) are present For example, some Namibian localities such as Lüderitz (Hutchings *et al*. 2009) may provide similar feeding opportunities to those off WSA. For a better understanding of the population structure in the region, research effort and photographic and genetic data collection need to cover more seasons, and include more sites within WSA and the rest of West Africa, including farther offshore, similar to some of the long-term ocean-basin wide identification studies such as the SPLASH (Structure of Populations, Levels of Abundance, and Status of Humpbacks) program in the North Pacific (Calambokidis *et al.* 2008), or YoNAH (Year of the North Atlantic Humpback, Smith *et al.* 1999). The strategic deployment of satellite tags off WSA may offer a more short-term solution to elucidate potential migratory routes and locations of unknown breeding or feeding areas (e.g. Hauser *et al.* 2010; Zerbini *et al*. 2006) and would help inform whether the current IWC management units are relevant to humpback whales in this region.

ACKNOWLEDGEMENTS

Research work during the main research project off WSA was supported by the National Research Foundation (NRF), South Africa, under Grant Number 2047517. The following other organizations are thanked: Earthwatch Institute (funding), The Mazda Wildlife Fund (through the provision of a field vehicle), SASOL (through the donation of two four-stroke engines), PADI Project AWARE (UK) (funding), the South African Navy (access to the shore-based look-out), the Military Academy, University of Stellenbosch (accommodation) and Iziko South African Museum (office space and support). JB gratefully received financial support in the form of bursaries from the NRF, the Society for Marine Mammalogy, University of Pretoria, and the Wildlife Society of South Africa (Charles Astley Maberley Memorial bursary). The Namibian Dolphin Project is supported by NACOMA (Namibian Coastal Conservation and Management Project), the Nedbank Go Green Fund, Mohammed bin Zayed Species Conservation Fund, the British Ecological Society, the Rufford Small Grants Foundation and the Namibia Nature Foundation. JB and TJQC received funding from the International Whaling Commission (IWC) to conduct between-region matching. CP, IC, and HCR are grateful to the Sackler Institute for Comparative Genomics, American Museum of Natural History, in particular Dr George Amato, Dr Rob DeSalle, Matt Leslie and Jacqueline Ay-Ling Loo. The fieldwork would have been impossible without the enthusiastic assistance of numerous Earthwatch and other volunteers, to all of whom we owe a big debt of gratitude. The following photographers (from various affiliations) took or contributed images in WSA: Blake Abernethy, Meredith Thornton, Ingrid Peters, Desray Reeb, Shaun Dillon, Lisa Mansfield, Tilen Genov, Stephanie Plön, André du Randt, Mike Meÿer, Darell Anders, Sharon du Plessis, and Rob Tarr; and in Namibia: Mike Lloyd, Ute von Ludwiger, Orlanda Sardinha and Francois Visser (Levo Tours). The scanning protocol and photographic database structure was based on one developed by P.J. Ersts (later modified by S. Cerchio and TJQC, WCS) as part of an regional Atlantic/Indian Ocean humpback whale research collaboration. South African digital coastline data were obtained from SANBI (South African National Biodiversity Institute) as used for the National Spatial Biodiversity Assessment (NSBA) 2004. Ocean-wide bathymetry shape files were downloaded from Natural Earth (http://www.naturalearthdata.com/). Steve Kirkman, Ingrid Peters and Salvatore Cerchio are thanked for discussions on mark-recapture analyses. An earlier version of the manuscript benefitted greatly from the comments of two anonymous referees. Research in South Africa was authorised by successive annual permits issued to PBB by the Minister for Environmental Affairs, in terms of Regulation 58 of the Marine Living Resources Act, 1998 (Act no. 18 of 1998). Research in Namibia was completed under a permit issued to SE & RL by the Namibian Ministry of Fisheries & Marine Resources.

REFERENCES

Allen JM, Carlson C, Viechnicki J and Stevick PT. 2008. Interim report: Research Contract 16, Antarctic Humpback Whale Catalogue. *Unpublished report.* Paper SC/60/SH19 presented at the IWC meeting held in Santiago, Chilé.

Arnason AN, Schwarz CJ.1999. Using POPAN-5 to analyse banding data. *Bird Study* 46 (supplement): S157-S168.

Barendse J, Best PB, Thornton M, Pomilla C, Carvalho I, Rosenbaum HC. 2010. Migration redefined? Seasonality, movements and group composition of humpback whales *Megaptera novaeangliae* off the west coast of South Africa. *African Journal of Marine Science* 32(1): 1-22.

Barlow J, Calambokidis J, Falcone EA, Baker CS, Burdin AM, Clapham PJ, Ford JKB, Gabriele CM, LeDuc R, Mattila DK, Quin II TJ, Rojas-Bracho L, Straley JM, Taylor BL, Urbán R, J, Wade P, Weller D, Witteveen BH, Yamaguchi M. 2011. Humpback whale abundance in the North Pacific estimated by photographic capture-recapture with bias correction from simulation studies. *Marine Mammal Science.* DOI: 10.1111/j.1748-7692.2010.00444.x

Best PB. 1994. A review of the catch statistics for modern whaling in Southern Africa, 1908–1930. *Reports of the International Whaling Commission* 44: 467–485.

Best PB, Allison C. 2010. Catch history, seasonal and temporal trends in the migrations of humpback whales along the west coast of southern Africa. Paper SC/62/SH5 presented at the IWC Scientific Committee in Agadir, Morocco.

Best PB, Sekiguchi K, Findlay KP. 1995. A suspended migration of humpback whales *Megaptera novaeangliae* on the west coast of South Africa. *Marine Ecology Progress Series* 118: 1–12.

Best PB, Reeb D, Morais M, Baird A. 1999. A preliminary investigation of humpback whales off Northern Angola. *Unpublished report.* Paper SC/51/CAWS33 presented at the IWC Scientific Committee meeting held in Grenada.

Blackmer AL, Anderson SK, Weinrich MT. 2000. Temporal variability in features used to photo-identify humpback whales (*Megaptera novaeangliae*). *Marine Mammal Science* 16(2): 338-354.

Bryden MM. 1985. Studies of humpback whales (*Megaptera novaeangliae*), Area V. In: Ling JK, Bryden MM (eds.). Studies of sea mammals in south latitudes. South Australian Museum, Adelaide, pp115-123.

Burnham KP, Anderson DR, White GC, Brownie C. Pollock KH. 1987. Design and analysis methods for fish survival experiments based on release–recapture. *Monograph 5, American Fisheries Society*, Bethesda, MD.

Calambokidis J, Barlow J. 2004. Abundance of blue and humpback whales in the eastern North Pacific estimated by capture-recapture and line-transect methods. *Marine Mammal Science* 20(1): 63-85.

Calambokidis JEA, Falcone TJ, Quinn TJ, Burdin AM, Clapham PJ, Ford JKB, Gabriele CM, LeDuc R, Mattila D, Rojas-Brach L, Straley JM, Taylor BL, Urbän R J, Weller D, Witteveen BH, Yamaguchi M, Bendlin A, Camacho D, Flynn K, Havron A, Huggins J, Maloney N. 2008. *SPLASH: Structure of populations, levels of abundance and status of humpback whales in the North Pacific.* Final report for contract AB133F-03-RP-00078. U.S. Department of Commerce, Western Administrative Center, Seattle, WA. 57 pp. Available at www.cascadiaresearch.org [Accessed 10 August 2011].

Calambokidis J, Steiger GH, Straley JM, Herman LM, Cerchio S, Salden DR, Urbán-Ramirez J, Jacobsen JK, Von Ziegesar O, Balcomb KC, Gabriele CM, Dahlheim ME, Uchida S, Ellis G, Miyamura Y, De Guevara-P PL, Yamaguchi M, Sato F, Mizroch SA, Schlender L, Rasmussen K, Barlow J, Quin II TJ. 2001. Movements and population structure of humpback whales in the North Pacific. *Marine Mammal Science* 17(4): 769-794.

Carothers AD. 1973. Capture recapture methods applied to a population with known parameters. *Journal of Animal Ecology* 42:125–146.

Carvalho I, Loo J, Collins TJQ, Pomilla C, Barendse J, Best PB, Hersch R, Leslie MS, Thornton M, Rosenbaum HC. 2010. Temporal patterns of population structure of humpback whales in West coast of Africa (B stock). *Unpublished report.* Paper SC/62/SH8 presented at the IWC Scientific Committee meeting held in Agadir, Morocco.

Cerchio S. 1998. Estimates of humpback whale abundance off Kauiai, 1989 to 1993: evaluating biases associated with sampling the Hawaiian Islands breeding assemblage. *Marine Ecology Progress Series* 175: 23-34.

Chittleborough RG. 1965. Dynamics of two populations of the humpback whale, *Megaptera novaeangliae* (Borowski). *Australian Journal of Marine and Freshwater Research* 16: 33-128.

Clapham PJ, Leimkuhler E, Gray BK, Matilla DK. 1995. Do humpback whales exhibit lateralized behaviour? *Animal Behaviour* 50: 73-82.

Collins TJQ, Cerchio S, Pomilla C, Loo J, Carvalho I, Ngouessono S, Rosenbaum HC. 2008. Revised estimates of abundance for humpback whale breeding stock B1: Gabon. *Unpublished report.* Paper SC/60/SH28 presented at the IWC Scientific Committee meeting held in Santiago, Chilé.

Cooch E, White G. 2006. *Program MARK: A gentle introduction* (9th edn). 53 pp. Available at http://www.phidot.org/software/mark/docs/book [accessed 5 February 2011].

Dawbin WH. 1966. The seasonal migratory cycle of humpback whales. In: Norris KS (ed.) *Whales, dolphins and porpoises.* University of California Press, Berkeley, pp. 146-196.

Elwen SH, Reeb D, Thornton M, Best PB. 2009. A population estimate of Heaviside's dolphins, *Cephalorhynchus heavisidii*, at the southern end of their range. *Marine Mammal Science* 25(1): 107-124.

Findlay KP, Best PB. 1995. Summer incidence of humpback whales on the west coast of South Africa. *South African Journal of Marine Science* 15: 279–282.

Findlay KP, Best PB. 1996. Humpback whales observed migrating past Cape Vidal, South Africa, 1988 – 1991. *Marine Mammal Science* 12(3): 354-370.

Friday NA, Smith TD, Stevick PT, Allen J. 2000. Measurement of photographic quality and individual distinctiveness for the photographic identification of humpback whales, *Megaptera novaeangliae*. *Marine Mammal Science* 16:355–374.

Friday NA, Smith TD, Stevick PT, Allen J, Fernald T. 2008. Balancing bias and precision in capture-recapture estimates of abundance. *Marine Mammal Science* 24(2): 253-275.

Gill PC, Burton CLK, Bannister J. 1995. Photographic resight of a humpback whale between Australia and Antarctic Area IV. *Marine Mammal Science* 11(1): 96-99.

Gowans S, Whitehead H. 2001. Photographic identification of northern bottlenose whales (*Hyperoodon ampullatus*): sources of heterogeneity from natural marks. *Marine Mammal Science* 17(1): 76-93.

Gubili C, Johnson R, Gennari E, Oosthuizen WH, Kotze D, Meÿer M, Sims DW, Jones CS, Noble LR. 2009. Concordance of genetic and fin photo identification in the great white shark, *Carcharodon carcharias*, off Mossel Bay, South Africa. *Marine Biology* 156: 2199-2207.

Gunnlaughson T, Sigurjónsson J. 1988. A note on the problem of false positives in the use of natural marking data for abundance estimation. *Reports of the International Whaling Commission (Special Issue)* 12: 143-145.

Hammond PS. 1986. Estimating the size of naturally marked whale populations using capture–recapture techniques. *Report of the International Whaling Commission (Special Issue) 8:* 253–282.

Hammond PS, Mizroch SA, Donovan GP. 1990. Individual recognition of cetaceans: Use of photo-identification and other techniques to estimate population parameters. *Report of the International Whaling Commission* (Special Issue 12): 439pp.

Hauser N, Zerbini AN, Geyer Y, Heide-Jørgensen MP, Clapham PJ. 2010. Movements of satellite-monitored humpback whales, *Megaptera novaeangliae*, from the Cook Islands. *Marine Mammal Science* 26(3): 679-685.

Herman LM, Pack AA, Rose K, Craig A, Herman EYK, Hakala S, Milette A. 2010. Resightings of humpback whales in Hawaiian waters over spans of 10–32 years: Site fidelity, sex ratios, calving rates, female demographics, and the dynamics of social and behavioral roles of individuals. *Marine Mammal Science.* DOI: 10.1111/j.1748-7692.2010.00441.x

Hutchings L, van der Lingen CD, Shannon LJ, Crawford RJM, Verheye HMS, Bartholomae CH, van der Plas AK, Louw D, Kreiner A, Ostrowski M, Fidel Q, Barlow RG, Lamont T, Coetzee J, Shillington F, Veitch J, Currie JC, Monteiro PMS. 2009. The Benguela Current: An ecosystem of four components. *Progress in Oceanography* 83(1-4): 15-32.

IWC (International Whaling Commission). 1998. Report of the Scientific Committee. Annex G. Report of the Sub-Committee on comprehensive assessment of Southern Hemisphere humpback whales. *Report of the International Whaling Commission* 48: 170–182.

IWC. 2010. Report of the Scientific Committee. Annex H. Report of the Sub-Committee on the other Southern Hemisphere whale stocks. Available at: http://iwcoffice.org/_documents/sci_com/SCRepFiles2010/Annex%20H%20FINAL-sq.pdf [Accessed on 7 February 2011].

IWS (International Whaling Statistics). 1964. International Whaling Statistics, Vol. LIII. *Det Norske Hvalråds Statistiske Publikasjoner*, Oslo, Norway. 77pp.

Jolly GM. 1965. Explicit estimates from capture-recapture data with both death and immigration – Stochastic model. *Biometrika* 52: 225-247.

Kaufman GD, Smultea MA, Forestell P. 1987. Use of lateral body pigmentation patterns for photographic identification of East Australian (Area V) humpback whales. *Cetus* 7 (1): 5-13.

Katona SK, Whitehead HP. 1981. Identifying humpback whales using their natural marking. *Polar Record* 20(128): 439-444.

Krützen M, Barré LM, Möller LM, Heithaus MR, Simms C, Sherwin WB. 2002. A biopsy system for small cetaceans: darting success and wound healing in *Tursiops* spp. *Marine Mammal Science* 18: 863–878.

Larsen F, Hammond PS. 2004. Distribution and abundance of West Greenland humpback whales (*Megaptera novaeangliae*). *Journal of Zoology, London.* 263: 343-358.

Lukacs PM, Burnham KP. 2005. Estimating population size from DNA-based closed capture-recapture data incorporating genotyping error. *Journal of Wildlife Management* 69(1): 396-403.

Meekan MG, Bradshaw CJA, Press M, McLean C, Richards A, Quasnichka S, Taylor JG. 2006. Population size and structure of whale sharks *Rhincodon typus* at Ningaloo Reef, Western Australia. *Marine Ecology Progress Series* 319: 275-285.

Mills LS, Citta JJ, Lair KP, Schwartz MK, Tallmon, DA. 2000. Estimating animal abundance using non invasive DNA sampling: promise and pitfalls. *Ecological Applications* 10(1): 283-294.

Mizroch SA, Beard JA, Lynde M. 1990. Computer assisted photo-identification of humpback whales. *Reports to the International Whaling Commission* (Special Issue 12): 63-70.

Mizroch SA, Herman LM, Straley JM, Glockner-Ferrari DA, Jurasz C, Darling J, Cerchio S, Gabriele S, Salden DR, Von Ziegesar O. 2004. Estimating the adult survival rate of central North Pacific humpback whales (*Megaptera novaeangliae*). *Journal of Mammology* 85: 963-972.

Olsen Ø. 1914. Hvaler og hvalfangst I Sydafrika. *Bergens Museums Aarbok 1914-15*. 5: 1-56 (Translated from Norwegian).

Palsbøll PJ, Allen JM, Bérubé M, Clapham PJ, Feddersen TP, Hammond PS, Hudson, RR, Jørgensen H, Katona SK, Larsen AH, Larsen F, Lien J, Mattila DK, Sigurjónsson J, Sears R, Smith T, Sponer R, Stevick P, Øien N. 1997. Genetic tagging of humpback whales. *Nature* 388: 767-769.

Perkins JS, Balcomb III KC, Nichols Jr. G, DeAvilla M. 1984. Abundance and distribution of humpback whales (*Megaptera novaeangliae*) in West Greenland waters. *Canadian Journal of Fisheries and Aquatic Science* 41: 533-536.

Perkins JS, Balcomb III KC, Nichols Jr. G, Hall AT, Smultea M, Thumser N. 1985. Status of the West Greenland humpback whale feeding aggregation, 1981-83. *Reports of the International Whaling Commission* 35: 379-383.

Pomilla C. 2005. Genetic structure of humpback whale (*Megaptera novaeangliae*) populations on Southern Hemisphere wintering grounds PhD Thesis, New York University, USA.

Pomilla C, Rosenbaum HC. 2006. Estimates of relatedness in groups of humpback whales (*Megaptera novaeangliae*) on two wintering grounds of the Southern Hemisphere. *Molecular Ecology* 15: 2541-2555.

Rasmussen K, Palacios DM, Calambokidis J, Saborío MT, Dalla Rosa L, Secchi ER, Steiger GH, Allen JM, Stone GS. 2007. Southern Hemisphere humpback whales wintering off Central America: insights from water temperature into the longest mammalian migration. *Biology Letters* 3(3):302-305.

Reisinger RR, Karczmarski L. 2010. Population size estimate of Indo-Pacific bottlenose dolphins in the Algoa Bay region, South Africa. *Marine Mammal Science* 26(1): 86-97.

Reisinger RR, de Bruyn PJN, Bester MN. 2011. Abundance estimates of killer whales at subantarctic Marion Island. *Aquatic Biology* 12: 177-185.

Rice M, Carlson C, Chu K, Dolphin W, Whitehead H. 1987. Are humpback whale population estimates being biased by sexual differences in fluking behaviour? *Report of the International Whaling Commission.* 37: 333-335.

Rosenbaum HC, Clapham PJ, Allen J, Nicole-Jenner M, Jenner C, Florez-Gonzaléz L, Urbän-Ramirez J, Paloma-Ladrón G, Mori K, Yamaguchi M, Baker CS. 1995. Geographic variation in ventral fluke pigmentation of humpback whale *Megaptera novaeangliae* populations worldwide. *Marine Ecology Progress Series* 124: 1-7.

Rosenbaum HC, Pomilla C, Mendez M, Leslie M, Best PB, Findlay KP, Minton G, Ersts PJ, Collins TJQ, Engel MH, Bonatto S, Kotze PGH, Meÿer M, Barendse J, Thornton M, Razafindrakoto Y, Ngouessono S, Vely M, Kiszka J. 2009. Population structure of humpback whales from their breeding areas in the South Atlantic and Indian Oceans. *PLoS ONE* 4: e7318. doi:10.1371/journal.pone.0007318.

Santos-Tieder L, Mizroch SA, Sims C. 2003. A new scanning protocol for black and white negatives. Unpublished protocol. National Marine Mammal Lab, National Oceanic and Atmospheric Administration (NOAA), USA. Cited in IWC 2004.

Scheidat M, Castro C, Gonzalez J, Williams R. 2004. Behavioural responses of humpback whales (*Megaptera novaeangliae*) to whale watching boats near Isla de la Plata, Machalilla National Park, Ecuador. *Journal of Cetacean Research and Management.* 6(1): 63-68.

Schwarz CJ, Arnason AN. 2006. Jolly-Seber models in MARK. In*:* Cooch E, White G (eds), *Program MARK: A gentle introduction* (9th edn). pp. 401-452. Available at http://www.phidot.org/software/mark/docs/book [accessed 5 February 2011].

Schwarz CJ, Seber GAF. 1999. A Review of Estimating Animal Abundance III. *Statistical Science* 14: 427-456.

Seber GAF. 1965. A note on the multiple-recapture census. *Biometrika* 52: 249-259.

Seber GAF. 1982. The estimation of animal abundance and related parameters (2nd edn). Wiley, New York.

Smith TD, Allen J, Clapham PJ, Hammond PS, Katona S, Larsen F, Lien J, Mattila D, Palsbøll PJ, Sigurjónsson J, Stevick PT, Øien N. 1999. An ocean-basin-wide mark-recapture study of the North Atlantic humpback whale (*Megaptera novaeangliae*). *Marine Mammal Science* 15: 1–32.

Stamation KA, Croft DB, Shaughnessy PD, Waples KA, Briggs SV. 2010. Behavioral responses of humpback whales (*Megaptera novaeangliae*) to whale-watching vessels on the southeastern coast of Australia. *Marine Mammal Science* 26(1): 98-122.

Stevick PT, Palsbøll PJ, Smith TD, Bravington MV, Hammond PS. 2001. Errors in identification using natural markings: rates, sources, and effects on capture-recapture estimates of abundance. *Canadian Journal of Fisheries and Aquatic Science* 58:1861-1870.

Stevick PT, Aguayo A, Allen J, Avila IC, Capella JJ, Castro C, Chater K, Rosa, LD, Engel MH, Félix F, Flórez-González L, Freitas A, Haase B, Llano M, Lodi L, Munoz E, Olavarría C, Secchi E, Scheidat M, Siciliano S. 2004. Migrations of individually identified humpback whales between the Antarctic Peninsula and South America. *Journal of Cetacean Research and Management* 6 (2): 109-113.

Straley JM, Quinn II TJ, Gabriele CM. 2009. Assessment of mark-recapture models to estimate the abundance of a humpback whale feeding aggregation in Southeast Alaska. *Journal of Biogeography.* 36(3): 427-438.

Wade PR, Kennedy A, LeDuc R, Barlow J, Carretta J, Shelden K, Perryman W, Pitman R, Robertson K, Rone B, Carlos Salinas J, Zerbini A, Brownell Jr. RL, Clapham PJ. 2011. The world's smallest whale population? *Biology Letters* 7: 83-85.

Waits JL, Leberg PL. 2000. Biases associated with population estimation using molecular tagging. *Animal Conservation* 3: 191-199.

Wedekin LL, Neves MC, Marcondes MCC, Baracho C, Rossi-Santos MR, Engel MH, Simões-Lopes PC. 2010. Site fidelity and movements of humpback whales (*Megaptera novaeangliae*) on the Brazilian breeding ground, southwestern Atlantic. *Marine Mammal Science* 26(4): 787–802.

White GC, Burnham KP. 1999. Program MARK: Survival Estimation from populations of marked animals. *Bird Study (Supplement)* 46: 120-138.

Wright JA, Barker RJ, Schofield MR, Frantz AC, Byrom AE, Gleeson DM. 2009. Incorporating genotype uncertainty into mark-recapture-type models for estimating abundance using DNA samples. *Biometrics* 65: 833-840.

Zerbini AN, Andriolo A, Heide-Jørgensen MP, Pizzorno JL, Maia YG, VanBlaricom GR, DeMaster DP, Simões-Lopes PC, Moreira S, Bethlem C. 2006. Satellite-monitored movements of humpback whales *Megaptera novaeangliae* in the Southwest Atlantic Ocean. *Marine Ecology Progress Series* 313: 295-304.

Zerbini AN, Clapham PJ, Wade PR. 2010. Assessing plausible rates of population growth in humpback whales from life-history data. *Marine Biology* 157: 1225-1236.

TABLES

Table 1. Photographic and genetic contributions to west South Africa (WSA) humpback whale database from various projects and sources. Total number of individuals identified by matching using combined identification features (including microsatellites). [* Indicates projects by the Mammal Research Institute (MRI); **These numbers include all images and biopsies collected and incorporated into the database. It does not take photo quality or matches into consideration].

Project description*	Study years	No. of images/biopsies collected**					Individuals identified
		Total	TF	RDF	LDF	Biop.	
Miscellaneous contributions	1983-2007	143	96	30	17	1	32
Cape Columbine humpback *	1993	104	30	37	37	6	9
West coast Heaviside's dolphin*	1997,1999-2001, 2008	98	19	33	46	13	18
Saldanha Bay humpback whale*	2001-2003	739	173	294	272	104	135
Saldanha Bay / St Helena Bay southern right whale*	2003-2007	736	192	300	244	92	95
	Entire database	1820	510	694	616	216	289

Table 2. Annual collection effort of photo identification and genetic data that contribute to the WSA humpback whale database, expressed as number of days on which at least one identification image or biopsy was collected or ‘collection days’. [* numbers in brackets indicate total days on which boat was deployed, when known; ‘x’ indicates months with no boat effort during dedicated MRI studies. Months within dashed outline indicate west coast Heaviside’s dolphin study period; Light-gray shading indicates dedicated humpback whale study at Saldanha Bay (with shore-based observations); dark-gray shading indicates boat-based study on southern right whales at St Helena Bay. Italicised numbers in different typeface in 2001-2007 show those months used for abundance estimates].

Year	Month*												Total days*
	Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug.	Sept.	Oct.	Nov.	Dec.	
1983		2											2
1984		1											1
1988	1							1					2
1989				1									1
1990	1										1		2
1992					1								1
1993										6(13)	1(5)		7 (18)
1997			1						1				2
1999		3(13)	1(13)										4 (26)
2000	0(4)	4(13)	1(16)	0(6)									5 (39)
2001	0(8)	0(14)	1(15)	1(7)			1(4)	4(11)	**4(14)**	**4(9)**	**3(9)**	**4(4)**	22 (95)
2002					1(7)	1(14)	4(8)	5(11)	**3(10)**	**5(14)**	**5(9)**	**2(9)**	26 (82)
2003	**7(9)**	**2(2)**							**1(2)**	**3(11)**	**3(12)**	**0(5)**	16 (41)
2004	**3(9)**								**2(8)**	**5(15)**	**4(9)**	**3(10)**	17 (51)
2005	**2(6)**	**1**							**2(9)**	**4(18)**	**3(18)**		12 (51)
2006									**0(1)**	**1(16)**	**8(17)**	**3(7)**	12 (41)
2007	**0(2)**	**0(7)**									**2**	**0(8)**	2 (9)
2008		1											1
All	14	14	4	2	2	1	5	10	13	28	30	12	135

Table 3. Details of sightings (date and position) of humpback whales involved in photographic matches between tail fluke catalogues of WSA (‘ZAW’) and Gabon (‘Gab’), and WSA and AHWC (sex determined from biopsies), and time elapsed between sequential resightings.

WSA ID (sex)	Date (and position) photographed off West South Africa	Matched to other catalogue ID	Date (and position) photographed in other region	Time between consecutive betwee catalogue resighting (and direction of movement)
TF-ZAW-03-017 (F)	2003/01/13 (33.013^{O}S, 17.774^{O}E and 33.064^{O}S, 17.825^{O}E); 2003/01/14 (32.702^{O}S,17.99^{O}E) 2008/02/05 (33.03^{O}S, 17.875^{O}E)	TF-Gab-03-124	2003/09/04 (9.264^{O}S, 1.928^{O}E)	234 d (N); 4.4 yr (S
TF-ZAW-04-005 (F)	2004/11/08 (32.665S, 17.988E)	TF-Gab-04-045	2004/09/26 (9.264^{O}S, 1.928^{O}E)	43 d (S)
TF-ZAW-01-005 (M)	2001/12/16 (33.021^{O}S, 17.86^{O}E); 2002/11/02 (33.005^{O}S, 17.849^{O}E); 2003/01/14 (33.031^{O}S, 17.825^{O}E and 32.674^{O}S, 17.877^{O}E); 2006/11/07 (32.674^{O}S, 17.935^{O}E)	TF-Gab-02-299	2002/08/06 (9.264^{O}S, 1.928^{O}E)	233 d (N); 88 d (S
TF-ZAW-05-007 (M)	2005/11/24 (32.551^{O}S, 18.026^{O}E)	ahwc3054	2005/12/22 (33.859^{O}S,18.278^{O}E)	28 d (S)
TF-ZAW-06-014 (M)	2004/12/01 (32.703^{O}S, 17.888^{O}E); 2006/11/22 (32.973^{O}S, 17.856^{O}E)	ahwc3055	2005/12/22 (33.859^{O}S,18.278^{O}E)	386 d (S); 11.2 mo (

Table 4. Minimum photographic and genetic individual catalogue size, by year and identification feature for WSA humpback whales. Correction for apparent survival (ϕ) is 0.96 from Mizroch *et al.* (2004); correction for matching error (*e*) is calculated false-negative rates for RDF/LDF (0.09 and 0.14), and for MS the mean allelic error of 0.065 (see text).

Year		Total no. individuals identified/year/feature (no. known from previous years)			
		TF	RDF	LDF	MS
1983		2	1	1	-
1984		2	-	-	-
1988		3	-	-	-
1989		1	3	-	-
1990		3	-	-	-
1992		2 (1)	2 (2)	1	-
1993		7	9	9	3
1997		1	2	3	-
1999		3	4 (1)	6 (1)	2
2000		2	5 (2)	7 (2)	5 (2)
2001		30 (3)	59 (3)	66 (5)	39 (3)
2002		33 (6)	61 (9)	51 (7)	38 (7)
2003		24 (7)	38 (8)	35 (10)	29 (7)
2004		20 (5)	26 (5)	25 (1)	27 (7)
2005		14 (0)	32 (4)	24 (2)	27 (5)
2006		24 (4)	31 (4)	32 (4)	22 (6)
2007		11 (1)	-	-	-
2008		1 (1)	3 (1)	3 (1)	1
Totals and (correction)	(none)	154	237	230	156
	(survival)	121	189	183	127
	(survival and error)	n/a	172	156	119

Table 5. Abundance estimates (N^*) for humpback whales at Saldanha Bay, WSA by the Chapman's modified Petersen method using separate identification features for one pair of capture periods (j_1 = Sept 2001 - Feb 2002; j_2 = Sept 2002 - Feb 2003). Photographs with quality and orientation ratings of 'poor' and lower were excluded from the analysis. Correction factor for dorsal fins refers to calculated false negative rates (RDF = 0.091 and LDF = 0.138, see Table 7), and for microsatellite the mean allelic error rate of 0.065, applied using the method of Stevick *et al.* 2001 (see text). Percentage bias calculated relative to the uncorrected estimator. [Legend: n/a = not applicable; n_1= no. of individuals identified during j_1 and n_2 during j_2; s_1 and s_2 = no. of sampling events in j_1 and j_2 respectively; m_2 = no. of individuals seen in j_1 and resighted in j_2; SE = standard error; CV = estimated coefficient of variation; LCI = lower 95% confidence intervals; UCI = upper 95% confidence intervals.]

Feature	Treatment	n_1	s_1	n_2	s_2	m_2	N^* ± SE	CV(N^*)	LCI	UCI	% bias
TF	uncorrected	15	15	16	18	3	67 ± 23.03	0.34	35	129	n/a
RDF	uncorrected	39	42	58	65	7	294 ± 81.77	0.28	172	502	
	corrected	38.70	-	57.30	-	7.70	265 ± 69.61	0.26	160	440	10
LDF	uncorrected	39	44	49	61	8	221 ± 56.02	0.25	136	361	
	corrected	38.20	-	47.08	-	9.28	182 ± 41.57	0.23	117	283	18
MS	uncorrected	34	37	41	51	9	146 ± 32.70	0.22	95	225	
	corrected	30.58	-	40.30	-	9.63	122 ± 25.28	0.21	81	182	17

Table 6. Selected parameter estimates and model selection criteria for three model variants in the POPAN version of Jolly-Seber open–population model in MARK 5.1, for different identification features. Photographs with quality and orientation ratings of ‘poor’ and lower were excluded from the analysis. The model estimates the super-population (*N*) for humpback whales that feed during spring/summer off west South Africa, as derived from capture-recapture data from six successive capture periods in 2001-2007 ($j_1 – j_6$, see Appendix I). Notations used: SE = standard error; CV = estimated coefficient of variation; LCI = lower 95% confidence intervals (UCI = upper 95% confidence intervals); $QAIC_c$ = Quasi-Akaike Information Criterion value; Mod. likel. = model likelihood; NP = number of parameters. Models are sorted according to ascending $QAIC_c$.

Model	*N* ± SE	CV(*N*)	LCI	UCI	$QAIC_c$	$\Delta QAIC_c$	Mod. likel.	NP
				Tail flukes				
2 - $\{\phi_{0.96}\,\beta_t\,p_t\}$	531 ± 346.7	0.65	192.22	1771.18	106.399	0	1	8
3 - $\{\phi_{0.96}\,\beta_t\,p_.\}$	301 ± 99.95	0.33	171.55	587.82	106.587	0.188	0.9103	4
1 - $\{\phi_t\,\beta_t\,p_t\}$	233 ± 112.7	0.48	116.14	620.00	112.112	5.7127	0.0575	11
				Microsatellites				
2 - $\{\phi_{0.96}\,\beta_t\,p_t\}$	528 ± 143.4	0.27	332.14	921.92	230.686	0	1	9
3 - $\{\phi_{0.96}\,\beta_t\,p_.\}$	400 ± 65.01	0.16	300.52	561.31	234.518	3.8323	0.1472	2
1 - $\{\phi_t\,\beta_t\,p_t\}$	496 ± 145.0	0.29	304.87	906.65	235.006	4.3198	0.1153	11
				Right dorsal fins				
2 - $\{\phi_{0.96}\,\beta_t\,p_t\}$	1035 ± 374.8	0.36	551.72	2116.89	198.517	0	1	7
1 - $\{\phi_t\,\beta_t\,p_t\}$	955 ± 495.1	0.52	419.87	2604.19	206.221	7.7041	0.0212	11
3 - $\{\phi_{0.96}\,\beta_t\,p_.\}$	681 ± 139.1	0.20	472.24	1032.05	231.399	32.882	0	3
				Left dorsal fins				
1 - $\{\phi_t\,\beta_t\,p_t\}$	1232 ± 773.8	0.63	454.61	3950.49	141.953	0	1	11
2 - $\{\phi_{0.96}\,\beta_t\,p_t\}$	1013 ± 497.8	0.49	449.43	2612.07	146.357	4.4036	0.1106	7
3 - $\{\phi_{0.96}\,\beta_t\,p_.\}$	760 ± 194.1	0.26	481.12	1269.69	178.347	36.394	0	2

Table 7. False negative rates (number missed matches as % of total number of matching opportunities) detected for humpback whale photographic identification features, west South Africa, using microsatellite matches as a control. Only pictures of quality and orientation > 'poor' were used (as for abundance estimates).

ID feature	Sample occasions	Matching opportunities	Confirmed matches	Missed matches	False negative rate (%)
MS (control)	88	32	-	-	-
LDF	58	29	25	4	13.8
RDF	49	22	20	2	9.09
TF	30	13	13	0	0

Table 8. Abundance estimates (N^*) for WSA humpback whales from the Chapman's modified Petersen estimator for various model configurations using double marked (TF plus alternative mark) humpback whales identified during first sampling period, and recaptures based on TF or alternative mark during second sampling period. SE = estimated standard error, CV = estimated coefficient of variation, LCI and UCI = lower and upper 95% confidence intervals. * An error correction of 0.065 for MS, 0.091 for RDF, and 0.138 for LDF was applied for n_2 and m_2 using method of Stevick *et al.* 2001.

Model	n_1	$n'_2{}^*$	$m'_2{}^*$	N^*	SE(N^*)	CV(N^*)	LCI	UCI
{n_1 = TF&RDF,n_2 = RDF, m_2 = RDF}	10	57.30	3.30	148	48.65	0.33	79	277
{n_1 = TF&RDF, n_2 = TF, m_2 = TF}	10	16	3	46	14.58	0.32	25	84
{n_1 = TF&LDF, n_2 = LDF, m_2 = LDF}	11	47.08	3.48	128	41.47	0.32	69	238
{n_1 = TF&LDF, n_2 = TF, m_2 = TF}	11	16	3	50	16.28	0.33	27	93
{n_1 = TF&MS, n_2 = MS, m_2 = MS}	9	40.30	2.14	131	51.50	0.39	62	275
{n_1 = TF&MS, n_2 = TF, m_2 = TF}	9	16	2	56	21.51	0.39	27	116

Figures

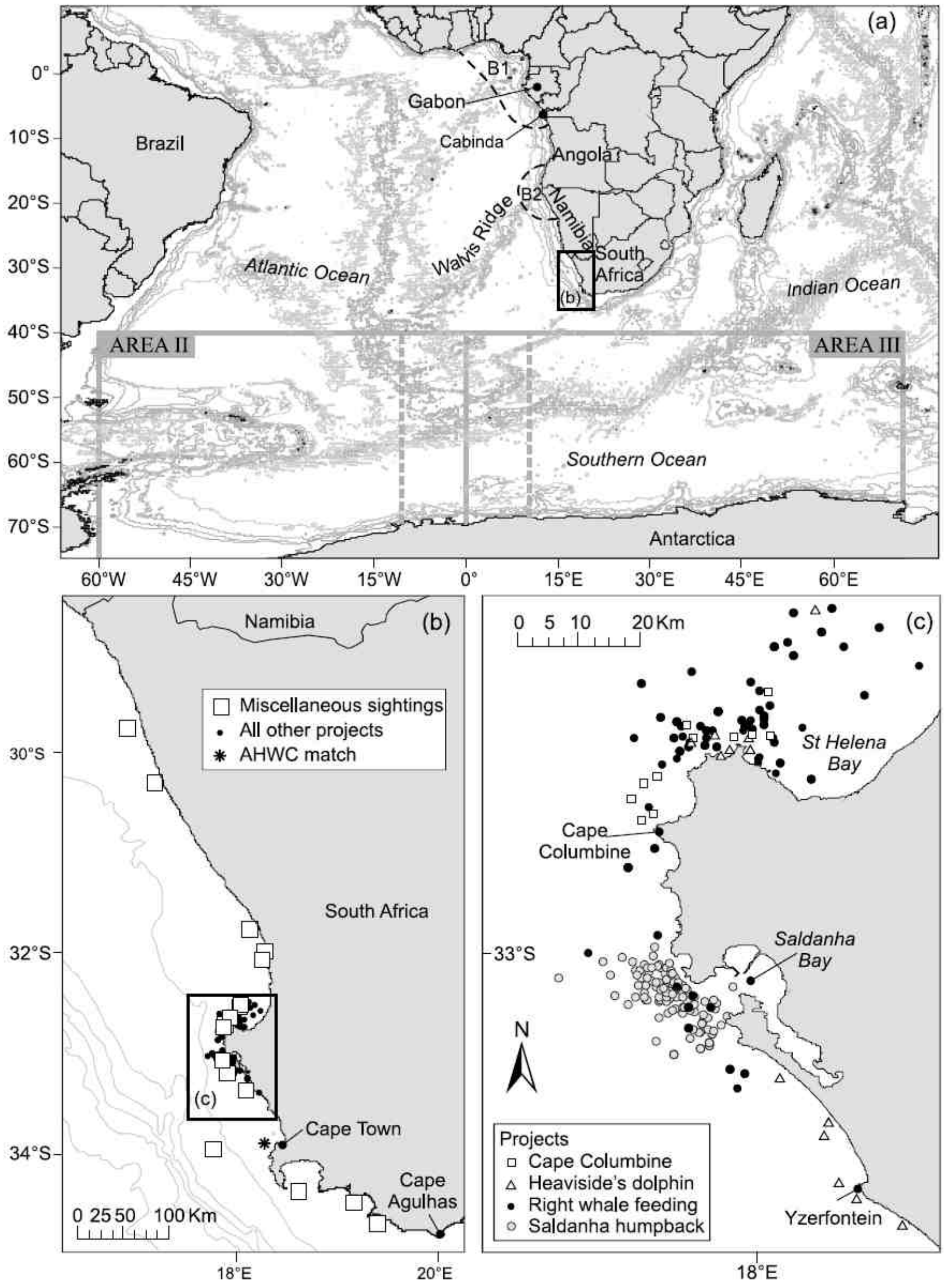

Figure 1: (a) The SE Atlantic-, SW Indian-and Southern Oceans showing bathymetry (to 4,000 m), areas of relevance to Breeding Stock B (BSB) southern hemisphere humpback whales, the speculated locations of sub-stocks B1/B2, Antarctic Feeding Areas II/III and suggested nucleus feeding area for BSB whales (10^{O}W – 10^{O}E, shown by dashed grey lines), and collection areas for regional photo-ID catalogues; (b) Detail of WSA and extent collection effort from various sources; (c) Detail of Saldanha/St Helena Bay area where the majority of data were collected during four major research projects, 1993 - 2007 (also see Tables 1 and 2).

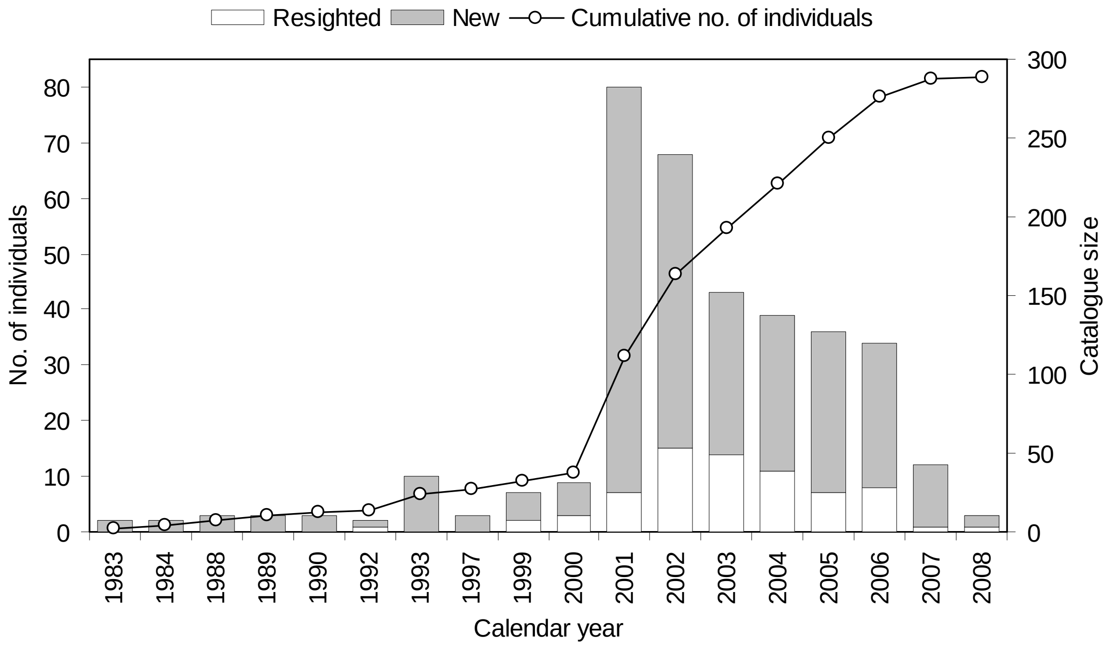

Figure 2. Total number of new and resighted individually identified humpback whales seen per year (1983 – 2008) off the west coast of South Africa, based on combined identification features (photographic and microsatellite), and cumulative number of individuals in the database (unadjusted for mortality).

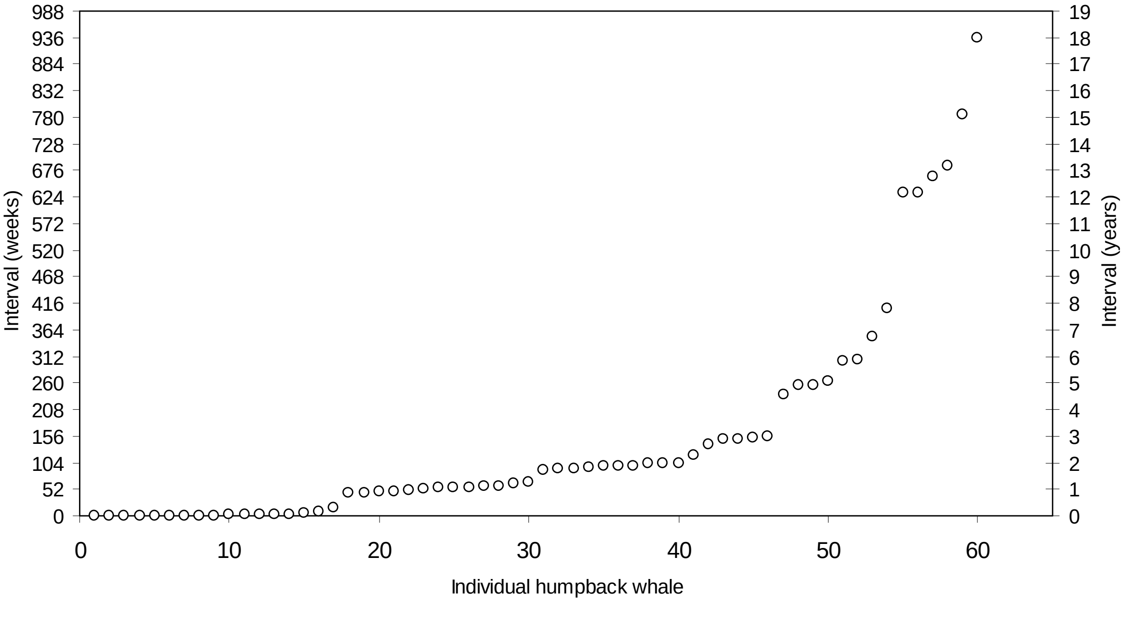

Figure 3. Intervals (in weeks and years) between the first and last sighting events for 60 humpback whales resighted on different days, off west South Africa.

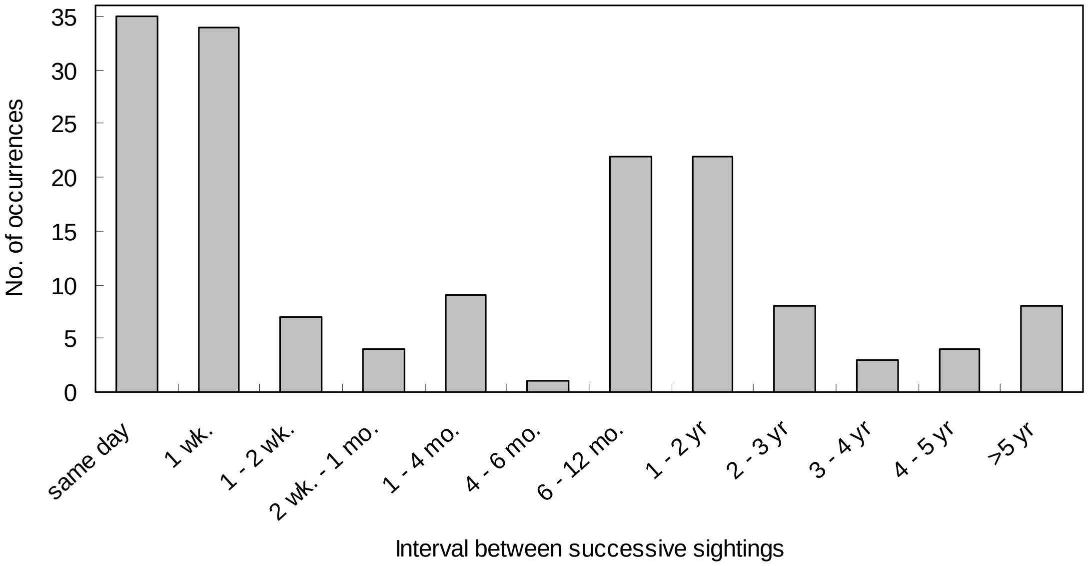

Figure 4. Frequency occurrence of time interval between successive sightings of 67 individually identified humpback whales off west South Africa (*n* = 157 sighting records).

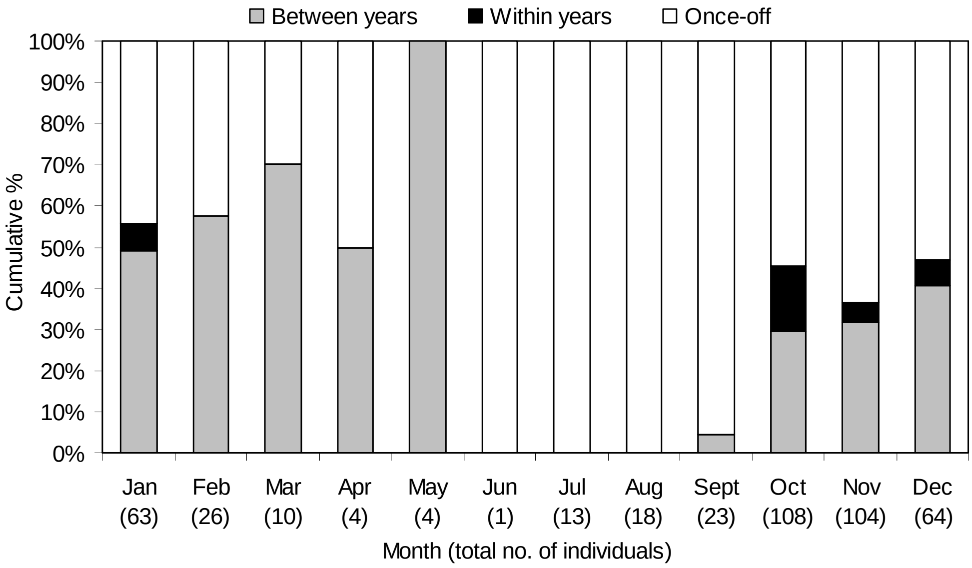

Figure 5. Time of year (month) off west South Africa during which humpback whales with different resighting histories (resighted within the same year only, resighted between years, and never resighted , i.e. once-off sightings) were recorded during 438 sighting events of 281 individual whales, based on combined identification features (1983 – 2008). Total number of unique individuals identified during a month is shown in brackets.

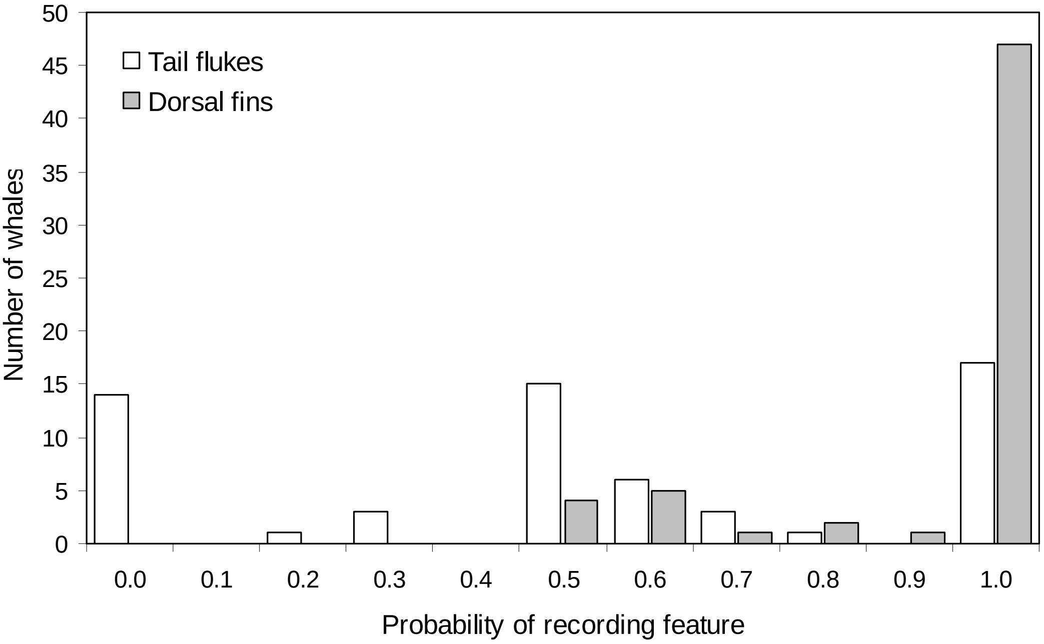

Figure 6. Probability for recording a photographic identification feature for 60 individual (resighted) humpback whales off WSA, calculated as the number of times a feature (tail fluke or dorsal fin) was recorded as proportion of the total number of times that the whale was intercepted.

APPENDIX I: SUMMARY CAPTURE-RECAPTURE STATISTICS

Summary capture-recapture statistics for individual identification features of humpback whales off west South Africa, and all features combined for six selected sampling periods, and used in CMP and JS abundance estimates. [j = sampling period; n = total whales identified per j; m = total recaptures per j; u = new identified whales; M = number of new whales before j]. Sampling Periods: j_1 = Sept 2001 - Feb 2002; j_2 = Sept 2002 - Feb 2003; j_3 = Sept 2003 - Feb 2004; j_4 = Sept 2004 - Feb 2005; j_5 = Sept 2005 - Feb 2006; j_6 = Sept 2006 - Feb 2007.

Tail flukes

m_j	j_1	j_2	j_3	j_4	j_5	j_6
j_1	x	3	1	0	0	0
j_2	-	x	0	1	0	1
j_3	-	-	x	1	0	0
j_4	-	-	-	x	0	0
j_5	-	-	-	-	x	1
M	0	3	1	2	0	2
N	15	16	10	7	9	16
U	15	13	9	5	9	14
M	0	15	28	37	42	51

Right dorsal fins

m_j	j_1	j_2	j_3	j_4	j_5	j_6
j_1	x	7	1	2	1	2
j_2	-	x	0	4	1	1
j_3	-	-	x	0	0	0
j_4	-	-	-	x	1	0
j_5	-	-	-	-	x	0
m	0	7	1	6	3	3
n	39	58	14	20	25	27
u	39	51	13	14	22	24
M	0	39	90	103	117	139

Left dorsal fins

m_j	j_1	j_2	j_3	j_4	j_5	j_6
j_1	x	8	1	0	1	0
j_2	-	x	0	0	0	1
j_3	-	-	x	0	0	0
j_4	-	-	-	x	0	1
j_5	-	-	-	-	x	1
m	0	8	1	0	1	3
n	39	49	11	16	13	28
u	39	41	10	16	12	25
M	0	39	80	90	106	118

Microsatellites

m_j	j_1	j_2	j_3	j_4	j_5	j_6
j_1	x	9	2	3	1	1
j_2	-	x	0	4	0	1
j_3	-	-	x	0	1	1
j_4	-	-	-	x	1	2
j_5	-	-	-	-	x	1
m	0	9	2	7	3	6
n	34	41	20	27	22	22
u	34	32	18	20	19	16
M	0	34	66	84	104	123