Photo identification of humpback whales *Megaptera novaeangliae* off West South Africa (Breeding Stock B2), and a preliminary (sub-) population estimate

JACO BARENDSE[1*], PETER B. BEST[1*], MEREDITH THORNTON[1], SIMON H. ELWEN[2], CRISTINA POMILLA [3, 4], INÊS CARVALHO [3, 4, 5] & HOWARD C. ROSENBAUM [4, 3]

[1]*Mammal Research Institute, University of Pretoria, c/o Iziko South African Museum, P.O. Box 61, Cape Town, 8000 South Africa.*
[2]*Mammal Research Institute, University of Pretoria, Pretoria, 0001 South Africa.*
[3] *Sackler Institute for Comparative Genomics, American Museum of Natural History, Central Park West at 79th street, New York, NY 10024, USA.*
[4]*Ocean Giants Program, Wildlife Conservation Society, 2300 Southern Blvd., Bronx, NY 10460-1099, USA.*
[5]*Faculdade de Ciências do Mar e Ambiente, Universidade do Algarve, Campus Gambelas, 8000-139 Faro, Portugal.*

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

ABSTRACT

An electronic image database was compiled for humpback whales photographed off the west coast of South Africa. The database incorporates all known sightings where photographs date back as far as 1983, up the end of February 2008. The final catalogue contains 510 pictures of tail flukes (TF), 616 of left dorsal fins (LDF), and 694 of right dorsal fins (RDF). Within and between-year matching was first carried out for each identification feature separately. Excluding images deemed ‘not useable’, this resulted in 154 different individuals being identified by TF, 230 by LDF, and 237 by RDF. Using combined features (TF, LDF, RDF, and microsatellite matches, based on 216 skin biopsies) a total of 289 individual whales were identified. Sixty-seven whales were seen more than once, including re-sightings on the same day or in the same year. Of these 44 were re-sighted in different years. The largest number of re-sightings for one individual was 11 times, seen in six different years. The longest interval recorded between the first and last events of identification was *ca.* 18 years for a whale first seen in 1989 and again in four subsequent years, the last being 2007. The re-sighting rate of 15.22% (based on combined identification features) at intervals of a year or more, appears to indicate a high level of fidelity to the region. More than 11% of whales were also seen on different days in the same year. The relatively low number of individuals identified by tail flukes, compared to dorsal fins, suggests that inconsistent fluking behaviour may introduce heterogeneity in sighting probabilities. Resightings between six different time-periods (spring and summer months in 2001-2007) were used to calculate preliminary abundance estimates for this sub-population, using closed population models.

ABUNDANCE ESTIMATES; BREEDING STOCK B2; HUMPBACK WHALE; MARK-RECAPTURE; MIGRATION; PHOTO-ID; SITE FIDELITY

INTRODUCTION

In the south-eastern Atlantic there remains a great deal of speculation around the exact breeding and feeding grounds utilised by humpback whales *Megaptera novaeangliae* from Breeding Stock B (BSB) (IWC 1998). This area, the west coast of Africa south of the equator, was characterised by extremely high catches from 1908 to 1914, and fluctuating catches thereafter (Best 1994). There is still much uncertainty regarding the relationship between whales found close inshore off the west coast of Southern Africa, and those that engage in breeding activities further north in the coastal waters of Gabon and Angola. It has been proposed that Region B consist of two sub-populations, B1 and B2, the former situated north of the Walvis Ridge or Angola/Benguela Front at about 18^0S and the latter south of this latitude (IWC 2001). Nuclear DNA analysis of samples from within Region B has not supported the sub-division of this region, though significant differences have been found between mitochondrial DNA haplotype frequencies of animals from west South Africa (B2), compared to areas further north (B1 - Gabon) (Carvalho *et al.* 2009, Rosenbaum *et al.* 2009). On the other hand, several microsatellite matches have now been recorded between whales biopsied off Gabon and west South Africa, providing the first direct evidence of individuals moving between these two areas (reported in Carvalho *et al.* 2009). The whales from BSB are thought to primarily migrate to Antarctic Areas II (60^0W to 0^0) and III (0^0 to 70^0E) for the summer. Althought the west coast of South Africa has traditionally been viewed as a migration corridor, presumably for B2 whales, some humpback whales here regularly display temporary residency during late spring and summer months, sometimes associated with feeding (Best *et al.*1995; Findlay & Best 1995; Barendse *et al.* in press), a behaviour also seen in southern right whales *Eubalaena australis* (Best *et al.* in prep.).

Prior to 1993 there was no directed research effort to investigate the population component of humpback whales that make use of this area (Best *et al*. 1995). We present here results from the most comprehensive photo-ID and genetic[1] collection effort to date off the west coast of South Africa in order to: (1) examine within and between-year resightings of individually identified whales using photographs of tail flukes, lateral views of dorsal fins, microsatellites, and all these features combined to establish residency rates; and (2) calculate (sub-) population estimates using the numbers of resightings or recaptures in mark-recapture models.

METHODS

Data collection

Boat intercepts

Images for inclusion in the west South Africa catalogue were obtained from two main sources: (1) those collected during two studies dedicated to humpback whales at Cape Columbine (1993) and Saldanha Bay (2001-2003), both of which included a shore-based watch component; and (2) from whales photographed when sighted incidentally during research work directed at other cetacean species, or seen during routine multi-disciplinary scientific cruises in the region, between 1983 and 2008 (Table 1 and Figure 1). At each intercept, the species was confirmed and GPS position and group size recorded. For all intercepts carried out by the Mammal Research Institute (MRI), attempts were made (for each individual) to take identification photographs of the ventral tail flukes and left and right lateral views of the dorsal fin and caudal peduncle. Prior to 2004, photographs were predominantly recorded on high speed (ISO400 and higher) black-and-white negative, colour positive and occasionally colour negative films, in most cases using motor-driven 35mm single lens reflex (SLR) cameras equipped with 100-300mm manual focus zoom lenses. Digital SLR's were used increasingly from 2004 up to January 2005 after which film cameras were no longer used.

Genetic sampling and analysis

During all boat intercepts carried out by the MRI attempts were made to collect a skin biopsy of each individual, using a Paxarms biopsy rifle (Krützen *et al.* 2002). Initially, some samples were lost due to the plastic darts cracking on impact at the thread holding the brass heads. The darts were modified (from 4 October 2001 onwards) by attaching a nylon monofilament tether to the main body of the dart and to the head to prevent it from falling off and sinking; this improved the retention of samples. All heads were decontaminated by flaming after use.

Skin samples were placed into individual cryogenic tubes filled with a NaCl-saturated, 20% dimethylsulfoxide (DMSO) solution. At the end of each day all skin samples were stored in a domestic freezer (-5°C) until they could be transferred to a -15°C freezer at the laboratory in Cape Town. Processing of samples was carried out by the Conservation Genetics Program at the American Museum of Natural History/Wildlife Conservation Society, New York.

Discrimination between individuals in the field (and association of specific images/biopsy attempts with individuals) was aided by onboard sketches of body features and recording of all photographic (film roll/data card numbers and frames) and biopsy sampling effort on each individual.

Days with collection effort ("collection days") were defined as those on which at least one picture or one biopsy was collected.

Photographic catalogue and sighting database

Management of images

Once processed (in the case of film) or downloaded (in the case of digital), photo frames or images were associated with specific individuals within specific groups on each day, using the field notes mentioned above. Film was scanned using either a film scanner (Canocraft FS2) or flatbed scanner with filmstrip adapter (EPSON SmartPanel or Canoscan FS8400). The scanning protocol and structure of the photographic database were based on those developed by P.J. Ersts (later modified by S. Cerchio, American Museum of Natural History) as part of an ongoing regional Atlantic/Indian Ocean humpback whale collaboration. Film frames were first scanned at 600 dpi with output dimensions of 9.843 cm (width) by 6.35 cm (height) and the scan window was scaled to maximise the coverage of the area of interest (i.e. tail fluke or dorsal fin). Black and white negatives were

[1] Genetic materials considered here are confined to samples collected by the Mammal Research Institute

scanned as colour film and later converted to 8-bit greyscale images (following the protocol developed by Santos Tieder *et al*. 2003). All these raw images were saved in the TIF format, after which no further size manipulations were carried out. For the purpose of importing images into the database copies of the raw images were converted to JPG format and reduced to resolutions of 100 and 200 dpi respectively for thumbnail and medium resolution copies. A similar protocol was applied to high resolution digital images, though the raw image format (JPG) was retained. Raw images were cropped to a height of 500 pixels and 200 dpi for medium resolution, and 250 pixels and 100 dpi for thumbnails.

Each image was individually assessed for photo quality, orientation of subject, and individual distinctiveness 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 TF image was further classified according to its ventral pigmentation pattern (or type) on a scale from 1-5, where 1 is all white (with no central black bar between the left and right flukes) and 5 all black (see Rosenbaum *et al*. 1995). Flukes were further rated for the part visible above water, viz. 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 obscuration of the subject (i.e. partial flukes, dorsal flukes or trailing/leading edges), or where the tail flukes were severely scarred or mutilated due to injury (e.g. killer whale bites).

Group and individual sighting histories

Daily sighting data for all groups of whales were entered into a Microsoft Access 2003 database. The minimum requirement for a group to be included in the database was that the sighting took place in South African waters west of Cape Agulhas (20^{O}E) and that the following information was available: the date (day, month and year) and locality (latitude and longitude). The boat name, photographer, sighting (group) number and group size (number of individuals) were recorded where possible and other information could include group composition and behaviour, SST and depth. An individual sighting incident (i.e. date, group nr. and individual designation) was entered for each individual of which at least one identification photograph (TF, LDF, RDF) was taken, or a biopsy collected, i.e. individual sighting incidents were not recorded for individuals with no photos taken or not biopsied. Images collected on the same date were compared to identify all individuals that were seen in more than one group on the same day (i.e. within-day matches). In the absence of sufficiently detailed group information, all sightings by the same boat on the same day were treated as a single sighting (this was only the case for some of the earlier years and for the few data obtained from non-MRI sources). Images were renamed, sorted into yearly folders, and imported into the database.

Photographic matching

Within years

Matching was done separately for each identification feature. Images were viewed on 15-19” TFT computer screens: thumbnail images first for the initial comparison, but when required, medium format and raw images to aid in the final decision making. Starting at the first year with data, within-year matching (checking for matches of the same individuals on different days in the same year) was carried out. For TF, the fluke types previously assigned were used to reduce the number of possible comparisons. Starting with TF of type 1, all images belonging to this type were compared to all other type 1’s. Furthermore, TF of a type were also compared to all images from the preceding and following types. For example, type 2 was compared to types 1, 2, and 3, and so forth. 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. A within-year identity number (ID) was assigned to each image (“Feature Year ID”), thus all available images for an individual whale seen in a year would be assigned the same within-year ID. Once all images were compared, a representative image (or images) for each identified individual was selected that would be used for between-year matching - these were assigned a between-year ID number (the “Feature Type ID”; the same as the “Feature Year ID”, but see below).

Between years

The selected images of each individual (with a unique between-year ID) from the earliest year of data were compared to those of the subsequent year in the database, in the same manner as described above. In the event of a positive match, the between-year ID of the later date was replaced by that of the earlier date. For example, an individual matched between 2000 and 2001 would have a unique within-year ID in each of these years but share the same between-year ID, which is the within-year ID assigned in the year it was first identified. Once all individuals between years were compared, one or more representative images were selected for each individual, all with the same between-year ID (“Feature Type ID”). These selected images would constitute the consolidated catalogue that included all identified individuals up to the end of the last year matched, including all preceding years. Following the same procedure the catalogue was compared to each subsequent year until the images of all unique between-year IDs had been compared to each other.

Verification of matches
The processes of within- and between-year matching were repeated by a second person for each identification feature. The within-year and between-year IDs were compared to identify any false-positives or negatives. These were reviewed and a decision made to accept or reject matches.

Microsatellite matching
The methodology for microsatellite genotyping is detailed in Carvalho *et al*. 2009. Each collected biopsy was associated to an individual sighting incident by its original biopsy number. In all cases where positive matches were made, the individual laboratory identity code assigned first (i.e. the earliest collected sample) was retained for that specific individual.

Combined identification features
Although matching was initially carried out for each feature independently, an individual sighting incident of a whale could contain up to four modes of identification viz. TF, RDF, LDF and microsatellite (MS). A sighting history (the complete collection of all sighting incidents captured in the database) was built using all matches made through all available identification modes between different sighting incidents. For example, a whale could be identified at the first incident by TF and RDF; it could then be matched to a second incident by TF only, during a sighting where a biopsy was also collected; finally, the whale could be matched by microsatellite to a third sighting where the whale was again biopsied. Although the whale was not identified by all modes at every incident, the complete sighting history contains a record of three modes of identification, collected across three sightings. An overall unique individual identification number (Catalogue ID) was assigned to each individual whale across its sighting history.

Mark-recapture abundance estimates

Individual distinctiveness and photo quality
The markings of some whales (on both tail flukes and dorsal fins) were more distinctive than others, for example due to prominent scarring. This could increase the likelihood of resighting such individuals, despite poor photo quality. In order to reduce the possible bias of less distinctive animals not being resighted when viewing photos of poor quality (i.e. increased false negatives; also see Stevick *et al*. 2001), all images with a quality and/or orientation rating of ‘poor’ and ‘not useable’, and a distinctiveness rating of ‘not useable’ were excluded from mark-recapture population estimates. In practise an individual distinctiveness of ‘not useable’ would invariably be a result of poor photo quality and orientation, as no individual could be judged outright to be ‘not useable’ simply because it was not very distinctive.

Data subset selection
Prior to 2001, collection of photo-ID data and genetic material tended to occur in a more or less *ad hoc* fashion (Table 1). To reduce the possible heterogeneity introduced by different seasonal attendance patterns of individuals, for the purpose of calculating population estimates, a subset of resighting data had to be selected, that was based on relatively consistent collection effort and seasonal coverage. Between 2001 and 2007 there was considerable research effort during spring (September, October, and November) and summer (December, January, and February) (Table 1), all by the MRI. Firstly there was a dedicated study (both shore- and boat-based) on humpback whales in the area at Saldanha Bay from May 2001 to February 2003 (33°02’S, 17°55’E). Secondly there was a boat-based study on southern right whale *Eubalaena australis* feeding at Saldanha Bay (in September) and St Helena Bay, about 30km further north (in October to December, rarely January) 2003 to 2006. Six ‘sampling periods’ (P) were therefore identified from 2001 - 2006, comprising the spring-summer season, starting in Sept 2001 to Feb 2002, Sept 2002 – Feb 2003, etc. up to Feb 2007. Although some months were not sampled during these periods (see Table 1 for details) it is assumed that sampling effort (i.e. number of collection days) during these subsets was consistent enough to allow calculations of preliminary abundance estimates of the population component that occurs in the region during the spring/summer feeding season (see Barendse *et al*. in press). Furthermore, because data were collected during successive seasons, pair-wise estimates could be calculated using closed population models. Also, with six different sampling periods, closed and open population models that consider multiple marking events could be attempted later, using this dataset.

Closed population models

The assumptions of closed models, relevant to using natural marks are (adapted from Seber 1982): (1) there is a constant population (*N*) during the sampling period (no immigration or emigration); (2) no marks are lost 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 (or 'marked') and newly sighted ('unmarked') whales have equal chance of being recaptured in subsequent samples.

Treating each pair of subsequent sampling periods (P) as a single mark-recapture event, the Chapman's modification to Petersen estimator (N^*) was calculated for resightings based on each separate identification feature, and combined features, using the formula (Seber 1982):

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

The variance (v^*) of N^* and the coefficient of variation (CV*) of N^* were calculated according to formulas in Seber 1982:

$$\hat{\mathrm{var}}(N^*) = v^* = \frac{(n_1+1)(n_2+1)(n_1-m2)(n_2-m2)}{(m_2+1)^2(m_2+2)} \quad (2)$$

And

$$\mathrm{CV}^* = \sqrt{\hat{\mathrm{var}}(N^*)} \Big/ N^* \quad (3)$$

The 95% confidence interval (CI) for the estimator N^* was calculated using a log-normal transformation as suggested by Burnham *et al*. (1987):

$$r = \exp\left(1.96\sqrt{\ln(1+(\mathrm{CV}(N^*))^2)}\right) \quad (4)$$

The upper CI was calculated by the product of N^* and *r,* while the lower was calculated by dividing N^* by *r*.

Although the sampled population could in reality never meet the conditions for a closed population, this estimator has been used on a number of occasions to calculate the size of feeding aggregations of humpback whales elsewhere on an inter-annual basis (e.g. Larsen and Hammond 2004, Straley *et al*. 2008), and is a reasonable approach for a long-lived mammal with relatively low rates of natural mortality and recruitment.

Due to the low number of resightings between pairs of sampling periods, the first three periods (P1-3) were combined to form a combined first sampling event (A), and the last three (P4-6) to form a second sampling event (B), and Chapman's modified Petersen estimate again calculated between these two events. Combining these periods may be in greater violation of, amongst others, the first (constant population) and the fifth assumptions (equal chance of being sighted during second period) due to mortalities and recruitment between the consecutive periods that make up the combined sampling events. Therefore, it was assumed that new ('unmarked') whales sighted in a sampling period had a 0.95 annual survival rate to the next (this rate is at the lower end of the range of survival estimates calculated for humpback whales; see Mizroch *et al*. 2004). Although crude, this is considered at least a partial attempt to correct for violations of the closed population model.

Other models

Raw capture-recapture data were provided to D. Butterworth and S. Holloway for further analysis, and are therefore not reported here.

RESULTS

Identification photos and sighting database

Collection effort

The greatest contribution of pictures (and number of identified individuals) was from the dedicated humpback whale study at Saldanha Bay from June 2001 to February 2003 (see Tables 1, 2 and Figure 1), which included both a shore-based watch and boat-based component. The second largest contribution was from a project examining southern right whale feeding in the years 2003 – 2007. This research was entirely boat-based and generally took place at Saldanha Bay in the month of September, and at St Helena Bay for the rest of the time. The only other notable contributions were from the 1993 humpback study at Cape Columbine (Best *et al*. 1995) and incidental humpback sightings during a project on Heaviside's dolphins *Cephalorhynchus heavisidii* (described in Elwen *et al*. 2009).

Photographic catalogue and sighting database

The west coast photo catalogue presently contains a total of 1,820 images, made up of 510 TF, 694 RDF, and 616 LDF images (Table 2). The sighting database includes records for 225 intercepts, and 447 sighting histories where images and/or biopsies were collected. Additions to the database from sightings prior to 2000 were few and contributed less than 10 individuals per annum (Figure 2). The biggest contribution was made in 2001 with the advent of the dedicated humpback whale study at Saldanha Bay, when a total of 80 whales were identified (seven were resightings from previous years). The additions to the catalogue remained at fairly high levels for the following five years, at over 25 individuals added per annum, although there was a steady decrease in the growth rate of the catalogue.

Matching

Individual ID-features

Excluding images that were deemed "not useable", 154 individuals were identified using only TF, 237 by RDF, and 230 by LDF. Microsatellite genotyping of the 216 skin biopsies yielded 38 samples matched to one or more other samples. This resulted in 156 individuals identified by this method. Three of these were identified by microsatellite only, i.e. were not photographed.

Combined ID-features

Using combined identification features, a total of 289 individual whales were identified, and excluding pictures classified as "not useable", 281 individuals remained. Images of eighteen tail flukes from different whales could not be linked to those of other features, or to biopsies, either due to the poor quality of the images, or to insufficient field-notes to match the various identification features. While these images (14 > "not useable") were used in the TF-only sighting histories (see above), they were not included in the combined feature catalogue, as these individuals were already represented by other identification features (dorsal fin pictures and/or biopsies) under unique Catalogue IDs. In this combined feature database, recorded during full sighting histories from 1983 to 2008, the majority of individuals were represented by all four identification features (Figure 3). This was followed by combinations containing both right and left dorsal fins. Sixty individuals were represented by single features only (3 MS, 17 LDF, 22 RDF, and 17 TF). The number of whales represented by TF, (135; alone and in combination with other features) was lower than that represented by all other methods: LDF (216), RDF (223) and MS (156) (Figure 3).

The discrepancy between the numbers of whales identified using single ID-features, and the number of individuals represented in the combined feature database, by the feature in question, is a result of missed matches (= false negatives), and in the case of TF, the 14 unassigned images mentioned above. The number of false negatives per feature relative to the combined features (i.e. identified by other means) could be calculated by the difference between total number of Feature Type IDs (for a single feature) and the total number of Catalogue IDs represented by the feature in question. This resulted in the detection of five missed TF matches (3.25%), 15 RDF (6.33%), and 14 LDF (6.09%) – detected false negatives as a percentage of total single-feature catalogue size.

Resightings

Out of the total number of individually identified whales, 222 were recorded in the database on one occasion only, and 67 more than once. Of the 55 whales that were resighted within years, 23 were only seen again in the same year, while the remaining 32 were also seen in other years. Forty-four whales were resighted between years and the number of between-year resightings ranged from one, to a maximum of six (Tables 3 and Figure 4). For both within and between-year resightings, 32 individuals where seen more than once as part of different groups on the same day. Furthermore, of the 44 between-year sightings, only 12 were not seen on multiple occasions in the same year (Tables 3 and 4), with one individual recorded 11 times (the same whale that was seen in six different years). The highest number of resightings between pairs of years (Table 4) was recorded between the years with the highest number of collection days, viz. 2001 - 2002 (14 resightings), 2002 - 2003 (10), and 2002 - 2003 (7).

Time between matches

The time (expressed in weeks) between the dates of first and last sightings was calculated for all individual whales that were resighted on different days (n = 60), 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 hours. Between-year time calculations took Leap years into account. The shortest time between resightings was one day (0.14 weeks), and the longest 934.86, or 18 years (based on 52 weeks per year). The average time between resightings was 178.13 weeks (3.4 years). For 14 whales, the time between the first and last sightings was longer than four years, while this interval was 12 years or longer for six of these (Figure 5).

Seasonality of resightings

Individual whale sightings were sorted by month, and separated on the basis of their resighting histories, viz. seen only once, resighted within years only, and resighted between different years. The latter may include within year sightings, but are not included in the "within-year only" category. All whales seen during winter months (June to August) were once-off sightings. During all other months a component of the whales seen was resighted on another occasion, the majority also between years. During October – January, a small proportion of resighted individuals were only resighted in the same year. However, from February – May all resighted individuals were also seen in other years, and the resighted component included 50% or more of all sightings made during these three months.

Abundance estimates

Chapman's modified Petersen estimator

The abundance estimates between successive sampling periods yielded variable results, in terms of their magnitude using different ID-features, and between different sampling periods (Table 5 and Figure 7). The estimates from TF resightings were the lowest overall, ranging between 28 and 93 (excluding the combined sampling events), and well below the minimum estimates from any other feature between any sampling periods (Table 5). The highest estimates in most instances were made from RDF resightings, even higher than for the combined features. Estimates based on microsatellite resightings were higher than for TF, but lower than both RDF and LDF for the first two periods. When estimates (based on the various identification features) were compared between the different sampling periods, there was a consistent trend of the P2-3 estimate being the highest and the P3-4 the lowest; the only exception to this was the result from the LDF dataset. The differences in the estimates between P1-2, and P2-3 varied greatly for different identification features. In contrast, the estimates for P3-4 (excepting LDF), P4-5 and P5-6 respectively, were mostly of the same order of magnitude, albeit much lower than those of the first two sampling periods, regardless of the ID feature used. By combining sampling periods, and adjusting for mortality between the successive periods, higher abundance estimate were obtained for all features, and the CV reduced. The LDF estimate again stands out as an anomaly, being higher than any other identification method. Also notable is that the estimates between the sampling events (A-B) did not differ greatly from the highest ones obtained for P2-3 for RDF, MS, and combined features.

Open population models

Due to time constraints no analysis using open population models was completed. Summaries of the recapture histories for individual ID-features and all combined are provided in Tables 6 (a) - (e).

DISCUSSION

Matching

Resightings

The high overall resighting rate of individuals (19% of all whales were seen more than once) confirms that the same individual whales return to the area. A further breakdown of the nature of these resightings highlights two major tendencies: Firstly, the prevalence of whales resighted on the same day, and those seen more than once in the same year (12.5%, ignoring same-day resightings) suggests that these whales are not merely moving through the area, as would be expected during a typical migration. Rather, they may move around locally, leaving some groups and joining others, and remain in the general area for days, weeks, and even months. Secondly, and perhaps more significant, is the high resighting rate of all whales matched between years (15.22%). The regular occurrence of annual or biennial resightings for the same whales, and the sometimes decade-long periods (recorded on six occasions), between the first and last recorded sightings, point to relatively long-term attendance patterns. These findings show that these whales routinely visit the coastal waters of this region, lending further support to historical accounts, and more recent behavioural evidence for the temporary residency of humpback whales. Furthermore, the seasonal (monthly) distribution of resightings suggests that humpback whales that engage in feeding during late spring, and in particular, summer months (as discussed in Barendse *et al.* in press) are also likely to be encountered repeatedly during the same season in other years.

Combined ID-features

The use of combined identification features may increase the ability to detect resightings of individuals and so build more comprehensive sighting histories; however, it also raises some concerns when these data are considered, for example, to calculate abundance estimates. Whales that were identified by more features earlier on are more likely to be resighted at later occasions. In some cases, at sightings where only a single feature was recorded that was never recorded again (in combination with another), this would result in a 'false negative', i.e. the incident with only one feature would have a unique Catalogue ID and resightings of this animal may go undetected. Stevick *et al*. 2001 concluded that while false positives are probably rare in photo-ID studies, false negatives are fairly common, with a higher error rate the poorer the quality of pictures. They developed a correction factor in order to reduce the positive bias (i.e. overestimation) resulting from false negatives (Stevick *et al*. 2001). Paradoxically, although the use of combined ID-features may result in false negatives in the combined feature database (and so positively bias the overall number of identified animals), it allowed the detection of (some) false negatives that occurred when using single ID-features. Dorsal fin matches appeared twice as likely to be missed compared to tail flukes (also see below).

Abundance estimates

ID-features

The variation between abundance estimates using different identification features, from an overall low of 28 (CI 95% 15, 55, CV = 0.35) for TF in P3-4, to a high of 442 (CI 95% 167, 1170, CV = 0.53) for RDF in P2-3, raises a number of points. Loss of marks is not considered a major issue with humpbacks, as both tail flukes and dorsal fins, along with the peduncle knobs are known to be very stable identification features (Blackmer *et al.* 2000).The remarkably low abundances calculated from TF resightings (even lower than the overall number of 154 whales identified by TF in the catalogue) for all sampling periods, are most likely a result of the small number of whales identified by this feature and possible heterogeneity of recapture probabilities The exposure of the ventral surfaces of the flukes is a behaviour known to vary between sexes (Rice *et al.* 1986), and fluking as an individual behavioural trait may reduce the overall number of whales identifiable by this feature. This was identified by Straley *et al.* 2008 as a factor that violates the assumption of equal capture probability for the whole population. The low proportion of whales represented by TF in the combined database (ca. 48%), compared to all other features (RDF = 79.4%, LDF = 76.9% and MS = 55.5%), may be testament to this, and may explain why TF estimates are much lower than those based on dorsal fins, which should theoretically always be exposed during surfacings. There may however be differences in the ability of researchers to obtain good quality images of these different features: during a typical approach from the rear, chances are good to obtain a TF picture (provided that they are shown). 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. Dorsal fins also have fewer distinguishing features (shape and peduncle knobs, when present) compared to tail flukes (shape, pigmentation, trailing edge), making it more difficult to match pictures of lower quality when the fin is not very distinctive. Errors as a result of the lack of distinctive markings definitely add heterogeneity to sampling probabilities (Stevick *et al.* 2001).

Estimates based on dorsal fin resightings were the highest obtained, especially during the first two sampling periods. While RDF estimates followed the general between-period trend, LDF deviated somewhat from this. Perhaps more surprising was the difference between RDF and LFD estimates, when both their single-feature catalogues are of almost equal size (237 and 230 respectively). Looking at the sub-sample selected for the abundance estimates, the low number of whales identified for P3, P4 and P5 using LDF (11, 16, and 13 respectively); the overall lower number of 'new' whales identified across P1-6 (sum of $M_{P1\text{-}6}$, RDF = 163, LDF = 143); and the low number of resightings (compared to RDF) using this feature, probably account for this difference. Also, recall that for abundance estimates, a quality 'filter' was applied that excluded all images classified as 'Poor', in an attempt to reduce heterogeneity caused by image quality and orientation. It would appear that it was more difficult to obtain LDF ID-photos of acceptable quality, compared to RDF. Shore-based tracking during late spring and early summer (during P1-3) showed that there were more southbound whales, i.e. with their left sides turned towards the shore (Barendse *et al.* in press); whether this influenced image quality may warrant further investigation.

Genotypic abundance estimates should be free from sampling heterogeneity applicable to photographs (i.e. image quality and fluking behaviour), and thus should better meet the assumptions of equal chances of capture during first sample, and recapture during the second sample (but see Mills *et al.* 2000 for discussion on negative biases applicable when using genotypic marks). As such the microsatellite (MS) abundance estimate may be considered independent from ones made using photo-ID (as suggested by Gubili *et al.* 2009), and is also not susceptible (to the same degree) to false-negatives. The MS estimates were not the highest obtained, but rather similar to those calculated using other ID-features. Notably, these estimates were higher than those obtained for TF, although the total number of individuals identified by these two ID-features are almost equal (TF = 154 and MS = 156). The number of resightings obtained using MS was much higher than for TF.

The use of combined ID-features to calculate abundance estimates may be questionable due to, amongst other reasons, the possible occurrence of false negatives (see above); however, the combination does maximize sample sizes (and minimize CVs) and the population estimates obtained appeared reasonably within the range of others calculated using single features. Using combined features may introduce some false negatives; on the other hand it may remove others that resulted from low picture quality and/or low distinctiveness.

Sampling periods

There were notable variations between estimates made for different sampling periods. The highest were obtained for the first two pair-wise estimates; generally higher than the actual number of individuals identified for the ID-feature in question. However, the numbers calculated for different ID-features for P1-2 and P2-3, were distributed across very wide ranges, for example, during P1-2 the RDF estimate was 294 (CI 95% 172, 502) compared to the TF estimate of 67 (CI 95% 35, 129), and CVs were high, especially for P2-3. The estimates obtained for the remaining pairs of sampling periods were considerably lower than for the first two (lower than actual individuals identified), and the variation between different ID modes was much reduced. Although the boat was available more or less the same number of days during all sampling periods, there may have been enough differences in the sampling strategy and collection effort to introduce a temporal heterogeneity. Throughout the first three sampling periods, all making up part of the humpback-directed study at Saldanha (except for Jan/Feb 2003), observers on the shore could spot whales at greater distances and assist the boat to locate humpback groups – this is reflected by the higher number of collection days, compared to the purely incidental intercepts of the latter three periods. By the pooling of sampling periods into two sampling events (A-B), the estimates were all elevated and CVs somewhat reduced (the LDF estimate is an exception to this). The numbers obtained compared well to the highest estimates obtained using single ID-features (again, excl. LDF), and the confidence intervals for the pooled samples of RDF, MS and combined features were smaller than for the pair-wise estimates. Another possible source of sampling heterogeneity between periods relates to camera equipment used: digital replaced film cameras during P3 and all images were in colour from P4 onwards. Whether digital photography produces superior pictures to film is still being debated (Markowitz *et al.* 2003, Mizroch 2003), however, it certainly increases the number of images taken per sighting, which should improve the chances of obtaining a good quality image.

Conclusion

The consolidated photo-ID and genotypic database for humpback whales, recorded in the west South Africa region, contributes to a better understanding of residency rates and long-term attendance patterns for this "component" of the humpback whale Breeding Stock B. The variability of abundance estimates between sampling periods is most likely a result of differences in data collection and sampling effort, and perhaps spatio-temporal effects, as the general trend between periods is reflected by all estimates, regardless of the ID-feature used. The considerable difference in the estimates obtained by using different ID-features, however, is of greater significance. Though the cause of this effect is not altogether clear, it does suggests that population estimates based on a single identification feature should be heeded with a degree of caution, especially when the ID-feature used may depend on an individually variable behaviour pattern (such as fluking) to be recorded. The calculated estimates presented here, although probably in violation of a number of closed-model assumptions, suggests that the (sub-) population of humpbacks that utilise this area for feeding during late spring and summer is small, probably not totalling more than 500 individuals. The relationship of these whales with the rest of BSB humpbacks and to those (perhaps strictly migratory) that occur here outside of the feeding period remains uncertain.

ACKNOWLEDGEMENTS

The authors wish to thank the following institutions for various forms of support: The South African National Research Foundation (NRF, funding), Earthwatch Institute (funding), The Mazda Wildlife Fund (through the provision of a field vehicle), SASOL (through the donation of two four-stroke engines), American Museum of Natural History (genetic analyses), PADI Project AWARE (UK) (funding), the South African Navy (access to the shore-based look-out), and the Military Academy, University of Stellenbosch (accommodation). 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).

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.

Blake Abernethy, Ingrid Peters, Desray Reeb, Shaun Dillon, Lisa Mansfield, Tilen Genov and Stephanie Plön all took pictures while with the MRI. André du Randt and Mike Meÿer (Marine and Coastal Management, South Africa) are thanked for contributing photographs from the west coast.

All work was carried out under 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).

The work for the main project was supported by the National Research Foundation, South Africa, under Grant number 2047517.

REFERENCES

Barendse, J., Best, P.B., Thornton, T., Pomilla, C., Carvalho, I. and Rosenbaum, H.C. (In Press). 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.

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

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

Blackmer, A.L., Anderson, S.K. and Weinrich, M.T. (2000). Temporal variability in features used to photo-identify humpback whales (*Megaptera novaeangliae*). *Marine Mammal Science* 16(2): 338-354.

Burnham, K.P., Anderson, D.R., White, G.C., C. Brownie, C. and Pollock, K.H. (1987). Design and analysis methods for fish survival experiments based on release–recapture. *Monograph 5, American Fisheries Society*, Bethesda, MD.

Carvalho, I., Loo, J., Pomilla, C., Leslie, M.C., Collins, T.J.Q., Barendse, J., Best, P.B. and Rosenbaum, H.C. (2009). Temporal patterns of population structure of humpback whales in west coast of Africa (B1-B2 sub-stocks) based on mitochondrial SNA and microsatellite variation. Unpublished report, IWC Scientific Committee. *Annual Report of the International Whaling Commission*. Report no. SC/61/SH6 presented at the IWC meeting held in Madeira, Portugal.

Elwen, S.H., Reeb, D., Thornton, M. and Best, P.B. (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, K.P., Best, P.B. (1995). Summer incidence of humpback whales on the west coast of South Africa. *South African Journal of Marine Science* 15: 279–282.

Gubili, C., Johnson, R., Gennari, E., Oosthuizen , W.H., Kotze, D., Meÿer, M., Sims, D.W., Jones, C.S. and Noble, L.R. (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.

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 (International Whaling Commission). (2001). Report of the Scientific Committee. *Journal of Cetacean Research and Management* 3: 26–27.

Krützen, M., Barré, L.M., Möller, L.M., Heithaus, M.R., Simms, C., Sherwin, W.B. (2002). A biopsy system for small cetaceans: darting success and wound healing in *Tursiops* spp. *Marine Mammal Science* 18: 863–878.

Larsen, F. and Hammond, P.S. (2004). Distribution and abundance of West Greenland humpback whales (*Megaptera novaeangliae*). *J. Zool. Lond.* 263: 343-358.

Markowitz, T.M., Harlin, A.D. and Würsig, B. (2003). Digital photography improves efficiency of individual dolphin identification. *Marine Mammal Science* 19:217–223.

Mills, L.S., Citta, J.J., Lair, K.P., Schwartz, M.K., and Tallmon, D.A. (2000). Estimating animal abundance using non invasive DNA sampling: promise and pitfalls. *Ecological Applications* 10(1): 283-294.

Mizroch, S.A. (2003). Digital photography improves efficiency of individual dolphin identification: A reply to Markowitz et al. *Marine Mammal Science* 19(1): 612-614.

Mizroch, S.A., Herman, L.M., Straley, J.M., Glockner-Ferrari, D.A., Jurasz, C., Darling, J., Cerchio, S., Gabriele, S., Salden, D.R. and Von Ziegesar, O. (2004). Estimating the adult survival rate of central North Pacific humpback whales (*Megaptera novaeangliae*). *Journal of Mammology* 85: 963-972.

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

Rice, M., Carlson, C., Chu, K., Dolphin, W. and Whitehead, H. (1987). Are humpback whale population estimates being biased by sexual differences in fluking behaviour? *Rep. Int. Whal. Commn.* 37: 333-335.

Rosenbaum, H.C., Clapham, P.J., Allen, J., Nicole-Jenner, M., Jenner, C., Florez-Gonzaléz, L., Jorge Urbän R., Paloma Ladrón, G., Mori, K., Yamaguchi, M. and Baker, C.S. (1995). Geographic variation in ventral fluke pigmentation of humpback whale *Megaptera novaeangliae* populations worldwide. *Marine Ecology Progress Series* 124: 1-7.

Rosenbaum, H.C., Pomilla, C., Mendez, M., Leslie, M., Best, P.B., Findlay, K.P., Minton, G., Ersts, P.J., Collins, T.J.Q., Engel, M.H., Bonatto, S., Kotze, P.G.H., 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, S.A. and Sims, C. (2003). A new scanning protocol for black and white negatives. Unpublished protocol. National Marine Mammal Lab, National Oceanic and Atmospheric Administration (NOAA). http://www.noaa.gov/

Seber, G.A.F. (1982). The estimation of animal abundance and related parameters, 2nd Ed. Willey, New York.

Stevick, P.T., Palsbøll, P.J., Smith, T.D., Bravington, M.V.and Hammond, P.S. (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.

Straley, J.M., Quinn II, T.J. and Gabriele, C.M. (2008). Assessment of mark-recapture models to estimate the abundance of a humpback whale feeding aggregation in Southeast Alaska. *J. Biogeogr.* Doi:10.111/j.1365-2699.2008.019606.x. 12pp.

TABLES

Table 1. Annual collection effort of photo identification (and genetic) data that contribute to the west South Africa humpback whale catalogue, expressed as number of days on which at least one identification image or biopsy was collected ("collection days")*.

YEAR	MONTH												Collection days
	J	F	M	A	M	J	J	A	S	O	N	D	
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
1997			1						1				2
1999	x	3(13)	1(13)	x							1		5
2000	0(4)	4(13)	1(16)	0(6)									5
2001	0(8)	0(14)	1(15)	1(7)	x	x	1(4)	4(11)	4(14)	4(9)	3(9)	4(4)	22
2002	x	x	x	x	1(7)	1(14)	4(8)	5(11)	3(10)	5(14)	5(9)	2(9)	26
2003	7(9)	2(2)	x	x	x	x	x	x	1(2)	3(11)	3(12)	0(5)	16
2004	3(9)	x	x	x	x	x	x	x	2(8)	5(15)	4(9)	3(10)	17
2005	2(6)	1	x	x	x	x	x	x	2(9)	4(18)	3(18)	x	12
2006	x	x	x	x	x	x	x	x	0(1)	1(16)	8(17)	3(7)	12
2007	0(2)	0(7)	x	x	x	x	x	x	x	x	2	0(8)	2
2008	x	1											1

*Numbers in brackets indicate total days on which boat was deployed; 'x' indicates months with no boat effort during chief MRI studies. Months within dashed outline indicate west coast Heaviside's dolphin study period; Light-gray shading indicates dedicated humpback study at Saldanha Bay (with shore-based observations); dark-gray shading indicates boat-based study on southern right whales at St Helena Bay. Months outlined in bold in 2001-2007 show those used for abundance estimates.

Table 2. Photographic and genetic contributions to west coast humpback whale database from various projects and sources. Asterisk indicates Mammal Research Institute (MRI) projects. Total number of individuals identified according to combined identification features (including microsatellites).

Project description	Study years	Nr 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**

**These numbers include all images and biopsies collected and incorporated into the database. It does not consider photo quality or matches

Table 3. Sighting histories of 44 identified humpback whales (based on combined identification features) that were resighted in different years. Numbers indicate total times resighted in the same year, and numbers with asterisk indicate resightings on the same day (e.g. 1, 2* indicates that the whale was seen three times in the same year, twice on the same day). Outlined cells show individual with longest recorded time, *ca.* 14 years, between 1st and last sightings). Photographs of quality rating 'Not useable' were not considered for resightings.

Indv. ID	Total nr of times seen	Resights between years	Year															
			1988	1989	1990	1992	1993	1997	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008
ZAW-043	2	2								1				1				
ZAW-069	2	2									1		1					
ZAW-075	2	2									1	1						
ZAW-091	2	2									1	1						
ZAW-235	2	2											1	1				
ZAW-269	2	2													1	1		
ZAW-286	2	2												1		1		
ZAW-292	2	2												1		1		
ZAW-295	2	2												1		1		
ZAW-011	3	3			1	1							1					
ZAW-028	3	2						1					2*					
ZAW-029	3	3						1			1	1						
ZAW-038	3	2							2		1							
ZAW-070	3	2									2		1					
ZAW-082	3	2									1	2*						
ZAW-085	3	2									1	2*						
ZAW-115	3	2									2*	1						
ZAW-118	3	2									1	2						
ZAW-126	3	3									1	1	1					
ZAW-170	3	2										1	2					
ZAW-173	3	2											1	2*				
ZAW-183	3	2										2	1					
ZAW-207	3	2											2		1			
ZAW-233	3	2											1	2*				
ZAW-273	3	2													2*	1		
ZAW-033	4	3							2			1		1				
ZAW-047	4	2									1		1, 2*					
ZAW-089	4	3									1	1	2*					
ZAW-097	4	2									1, 2*			1				
ZAW-107	4	3							1		1					2*		
ZAW-174	4	2									1	3*						
ZAW-204	4	2											3	1				
ZAW-210	4	2											3*	1				
ZAW-213	4	2											1, 2*					1
ZAW-240	4	2												3	1			
ZAW-096	5	4									1	1	2*			1		
ZAW-163	5	2										2, 2*			1			
ZAW-009	6	5		1							1	2*				1	1	
ZAW-019	6	5					1		1	2				1		1		
ZAW-036	6	5							1	1	1		1		2			
ZAW-015	8	4				1					2		2, 2*		1			
ZAW-017	8	3					4, 2*							1	1			
ZAW-101	8	4									1, 2*	2*	1	2*				
ZAW-006	11	6	1						2	2	1	2	1, 2*					

Table 4. Number of resightings of individual humpback whales (n=44) between pairs of calendar years, using combined identification features (TF, RDF, LDF, and microsatellites).

Year	89	90	92	93	97	99	00	01	02	03	04	05	06	07	08
1988	0	0	0	0	0	1	1	1	1	1	0	0	0	0	0
1989		0	0	0	0	0	0	1	1	0	0	0	1	1	0
1990	-		1	0	0	0	0	0	0	1	0	0	0	0	0
1992	-	-		0	0	0	0	1	0	2	0	1	0	0	0
1993	-	-	-		0	1	1	0	0	0	2	2	0	0	0
1997	-	-	-	-		0	0	1	1	1	0	0	0	0	0
1999	-	-	-	-	-		3	4	2	2	2	2	1	0	0
2000	-	-	-	-	-	-		2	1	2	2	2	0	0	0
2001	-	-	-	-	-	-	-		14	10	2	2	3	1	0
2002	-	-	-	-	-	-	-	-		7	2	1	2	1	0
2003	-	-	-	-	-	-	-	-	-		6	3	1	0	1
2004	-	-	-	-	-	-	-	-	-	-		3	3	0	0
2005	-	-	-	-	-	-	-	-	-	-	-		2	0	0
2006	-	-	-	-	-	-	-	-	-	-	-	-		1	0
2007	-	-	-	-	-	-	-	-	-	-	-	-	-		0

Table 5. Population estimates (*N) by the Chapman's modified Petersen method using separate identification features (TF, RDF, LDF and microsatellites), and these features in combination. Estimates were made between pairs of adjacent sampling periods (P1 - 6, see Table 1 for details), and between pooled sampling events (A = sampling periods 1 - 3, and B = 4 - 6). Photographs of quality rating 'Poor' were excluded from the analysis. The value of n_1 for the estimates between A - B has been adjusted for mortality between sampling periods, assuming a 0.95 survival rate per period.**

ID feature	P	n_1	n_2	m_2	N^*	CV(N^*)	SE(N^*)	lower CI	upper CI
Tail flukes	1-2	15	16	3	67	0.34	23.03	35	129
	2-3	16	10	1	93	0.50	45.87	37	232
	3-4	10	7	2	28	0.35	9.89	15	55
	4-5	7	9	0	79	0.64	50.20	25	247
	5-6	9	16	2	56	0.39	21.51	27	116
	A-B	**33**	**32**	**4**	**223**	**0.35**	**77.93**	**115**	**434**
Right dorsal fins	1-2	39	58	7	294	0.28	81.77	172	502
	2-3	58	14	1	442	0.53	233.77	167	1170
	3-4	14	20	6	44	0.22	9.49	29	67
	4-5	20	25	3	136	0.37	50.52	67	275
	5-6	25	27	3	181	0.38	69.32	88	374
	A-B	**92**	**72**	**12**	**521**	**0.23**	**117.36**	**337**	**806**
Left dorsal fins	1-2	39	49	8	221	0.25	56.02	136	361
	2-3	49	11	1	299	0.52	154.92	115	778
	3-4	11	16	0	203	0.66	133.99	62	660
	4-5	16	13	1	118	0.51	59.75	46	301
	5-6	13	28	3	101	0.35	35.62	51	197
	A-B	**80**	**57**	**4**	**939**	**0.38**	**355.19**	**458**	**1923**
Microsatellites	1-2	34	41	9	146	0.22	32.70	95	225
	2-3	41	19	2	279	0.45	124.38	121	643
	3-4	19	28	7	72	0.22	15.93	46	110
	4-5	28	22	3	166	0.38	62.93	81	340
	5-6	22	22	6	75	0.25	18.59	46	121
	A-B	**75**	**72**	**16**	**321**	**0.18**	**58.47**	**225**	**457**
Combined	1-2	58	64	16	225	0.17	38.55	161	314
	2-3	64	20	3	340	0.39	133.02	162	713
	3-4	20	34	10	66	0.17	11.02	48	91
	4-5	34	27	5	162	0.31	49.81	90	292
	5-6	27	33	7	118	0.25	29.32	73	191
	A-B	**109**	**94**	**22**	**453**	**0.16**	**71.80**	**333**	**617**

Table 6 (a) - (e). Summary capture-recapture statistics for individual identification features, and all features combined for six selected sampling periods. Notations used: P = sampling period; n = total whales identified per P; m = total resightings/P; u = new identified whales; M = number of new whales before P. (P1 = Sept 2001 - Feb 2002; P2 = Sept 2002 - Feb 2003; P3 = Sept 2003 - Feb 2004; P4 = Sept 2004 - Feb 2005; P5 = Sept 2005 - Feb 2006; P6 = Sept 2006 - Feb 2007).

(a) Tail flukes

m_P	P1	P2	P3	P4	P5	P6
P1		3	1	0	0	0
P2	-		0	1	0	1
P3	-	-		1	0	0
P4	-	-	-		0	0
P5	-	-	-	-		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

(b) Right dorsal fins

m_P	P1	P2	P3	P4	P5	P6
P1		7	1	2	1	2
P2	-		0	4	1	1
P3	-	-		0	0	0
P4	-	-	-		1	0
P5	-	-	-	-		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

(c) Left dorsal fins

m_P	P1	P2	P3	P4	P5	P6
P1		8	1	0	1	0
P2	-		0	0	0	1
P3	-	-		0	0	0
P4	-	-	-		0	1
P5	-	-	-	-		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

(d) Microsatellites

m_P	P1	P2	P3	P4	P5	P6
P1		9	2	3	1	1
P2	-		0	4	0	1
P3	-	-		0	1	1
P4	-	-	-		1	2
P5	-	-	-	-		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

(e) Combined features

m_P	P1	P2	P3	P4	P5	P6
P1		16	2	4	1	2
P2	-		1	5	1	0
P3	-	-		1	1	1
P4	-	-	-		2	2
P5	-	-	-	-		2
m	0	16	3	10	5	7
n	58	64	20	34	27	33
u	58	48	17	24	22	26
M	0	58	106	123	147	169

Figure 1. Map of study area and locations of humpback whale photo identification and genetic data collection effort during various research projects (see also Tables 1 and 2 for timing of effort and description of projects).

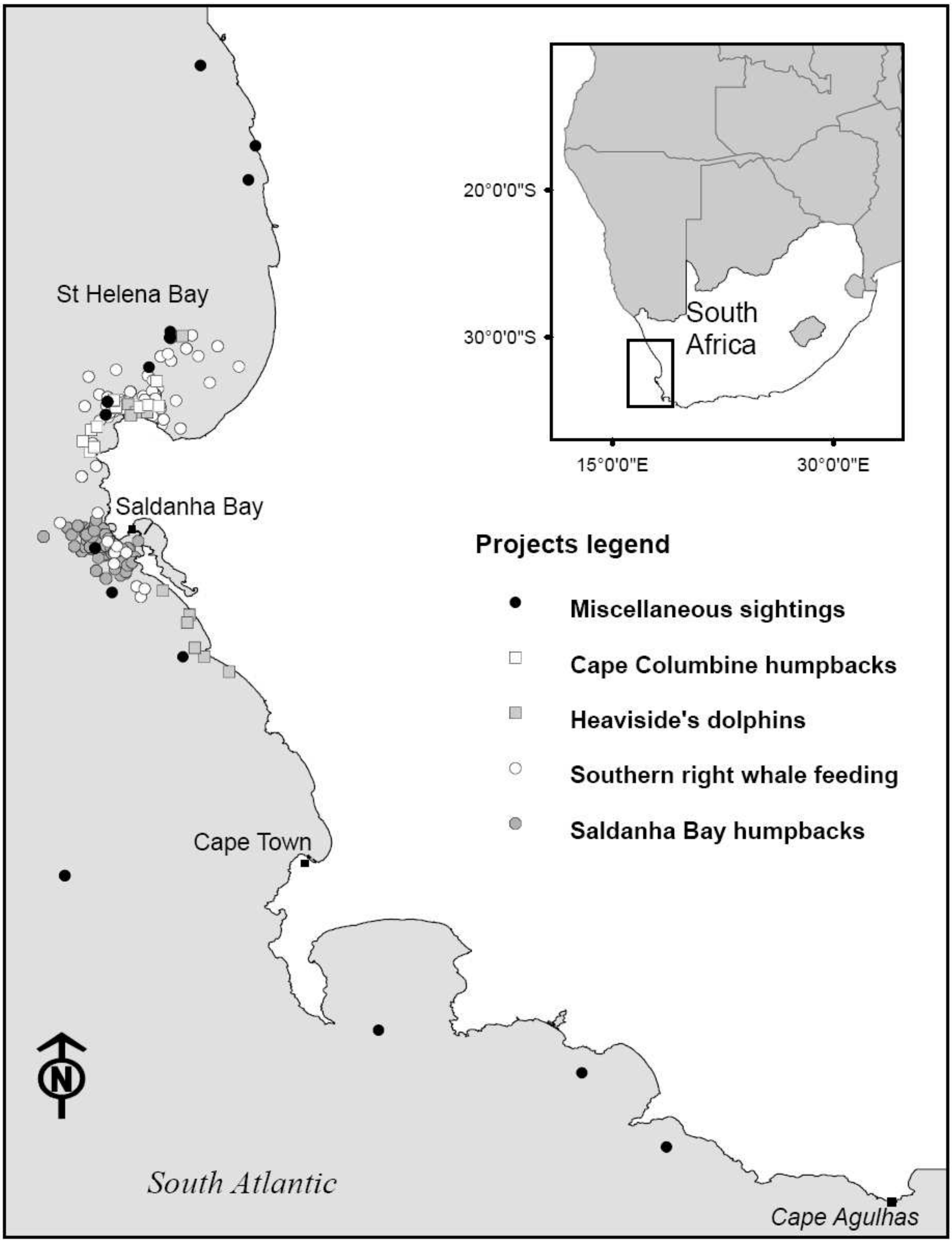

Figure 2. Yearly additions of individually identified humpback whales (using combined ID-features) to the west coast database, and cumulative growth (total number of unique individuals up to and including that year) of database between 1983 and 2008. Resighted individuals are those matched to preceding years.

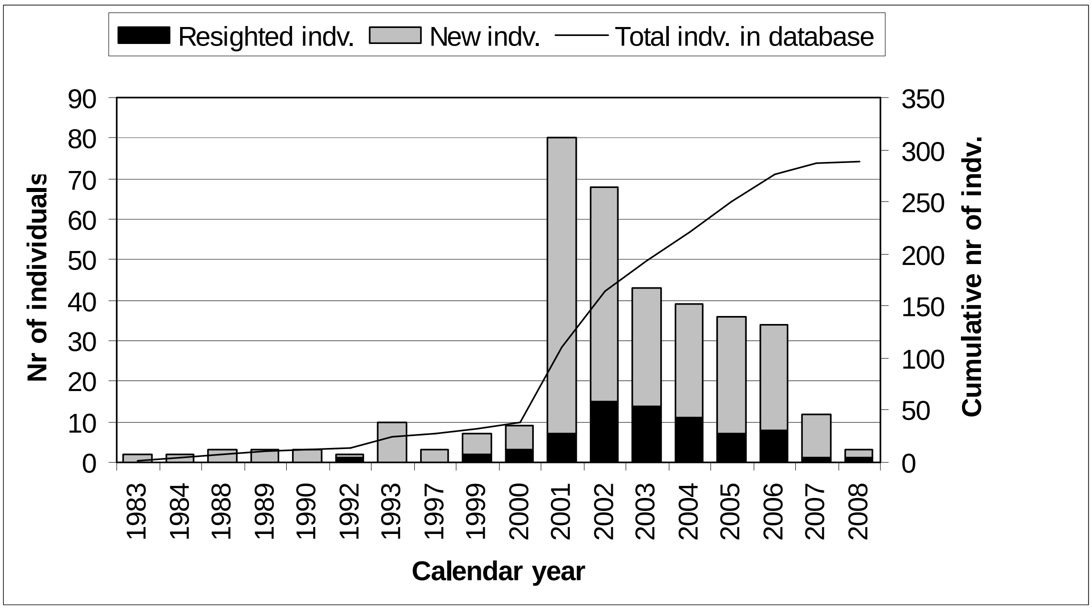

Figure 3. Number of humpback whales represented by various combinations of identification features in the west South Africa database (Key: TF = tail flukes, RDF = right dorsal fin, LDF = left dorsal fin, MS = microsatellite). Note that this is based on the full sighting histories of the whales between 1983 and 2008.

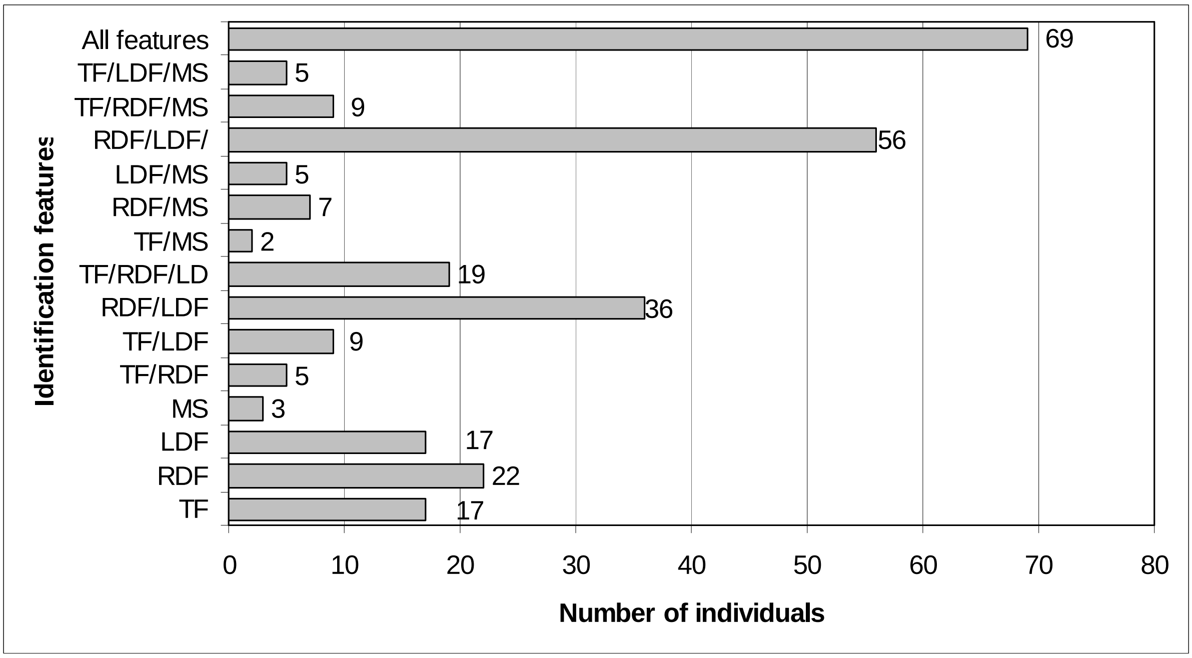

Figure 4. Number of resightings of individually identified humpback whales by category of resighting.

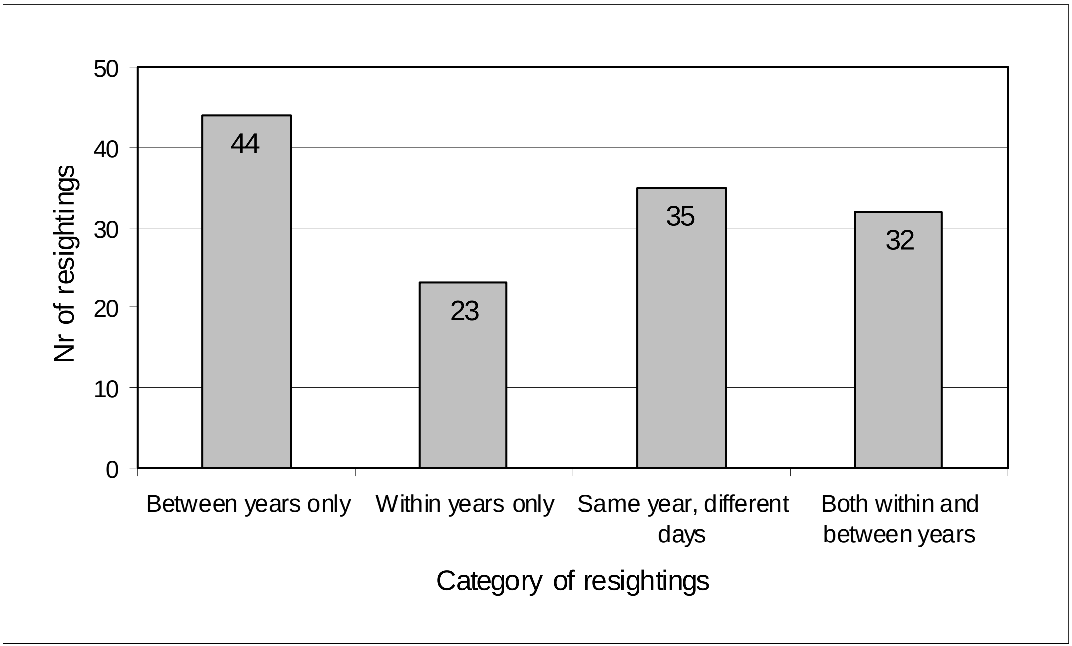

Figure 5. Time (in weeks) between first and last sighting events of individually identified humpback whales that were resighted on different days, within- and between different calendar years (n = 60).

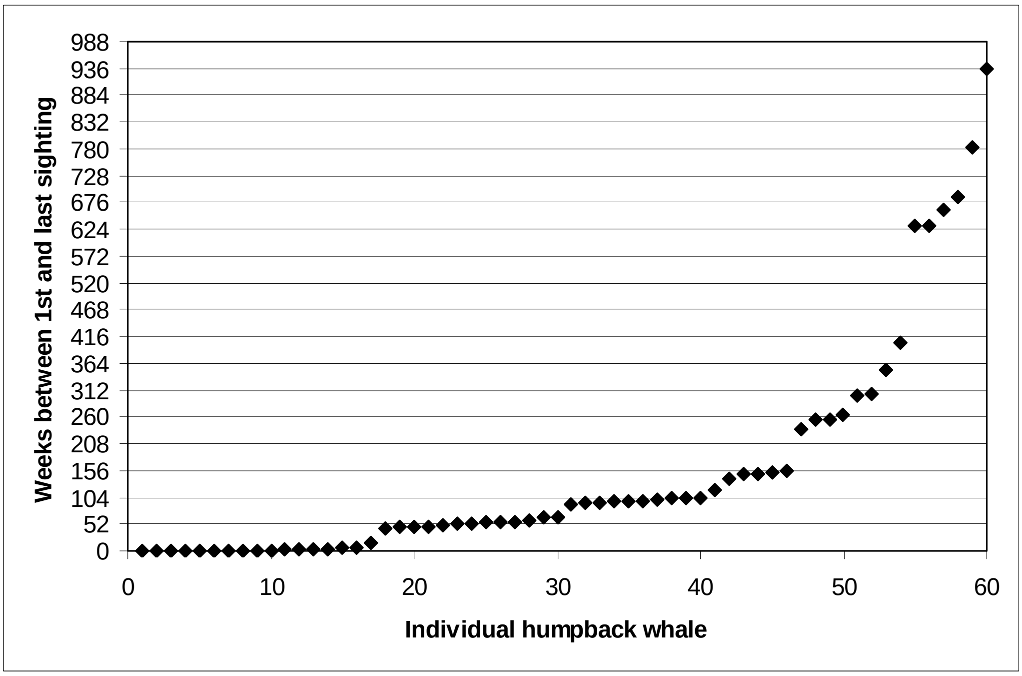

Figure 6. Monthly proportions of whales that were resighted within years only (on different days), between years (these may have been seen within-year), and not resighted (once-off sightings) between 1983 and 2008. Total number of identified individuals is shown in brackets.

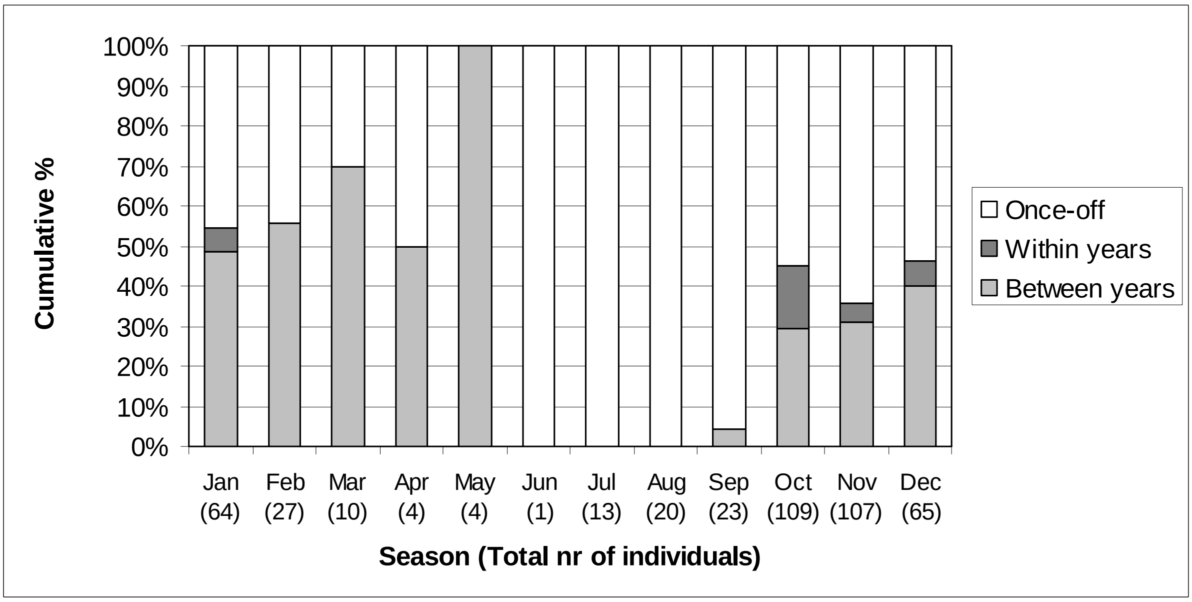

Figure 7. Abundance estimates (*N) calculated using the Chapman's modified Petersen estimator between successive sampling periods (P), and pooled sampling events (A-B), using individual identification features, and all features in combination (see also Table 5). (Key to legend: TF = tail flukes, RDF = right dorsal fins, LDF = left dorsal fins, MS = microsatellite).**