



Spotting the difference: towards fully-automated population monitoring of African penguins *Spheniscus demersus*

Richard B. Sherley^{1,2,*}, Tilo Burghardt^{3,4}, Peter J. Barham^{2,5}, Neill Campbell³,
Innes C. Cuthill¹

¹Centre for Behavioural Biology, School of Biological Sciences, University of Bristol, Woodland Road, Bristol BS8 1UG, UK

²Animal Demography Unit, Department of Zoology, University of Cape Town, Rondebosch 7701, South Africa

³Department of Computer Science, University of Bristol, Woodland Road, Bristol BS8 1UB, UK

⁴Computerised Monitoring and Biometric Identification in Natural Environments (COMBINE), Merchant Venturers Building, Woodland Road, Bristol BS8 1UB, UK

⁵H.H. Wills Physics Laboratory, University of Bristol, Tyndall Avenue, Bristol BS8 1TL, UK

ABSTRACT: Placing external monitoring devices onto seabirds can have deleterious effects on welfare and performance, and even the most benign marking and identification methods return sparse population data at a huge time and effort cost. Consequently, there is growing interest in methods that minimise disturbance but still allow robust population monitoring. We have developed a computer vision system that automatically creates a unique biometric identifier for individual adult African penguins *Spheniscus demersus* using natural markings in the chest plumage and matches this against a population database. We tested this non-invasive system in the field at Robben Island, South Africa. False individual identifications of detected penguins occurred in less than 1 in 10 000 comparisons ($n = 73\,600$, genuine acceptance rate = 96.7%) to known individuals. The monitoring capacity in the field was estimated to be above 13% of the birds that passed a camera ($n = 1453$). A significant increase in this lower bound was recorded under favourable conditions. We conclude that the system is suitable for population monitoring of this species: the demonstrated sensitivity is comparable to computer-aided animal biometric monitoring systems in the literature. A full deployment of the system would identify more penguins than is possible with a complete exploitation of the current levels of flipper banding at Robben Island. Our study illustrates the potential of fully-automated, non-invasive, complete population monitoring of wild animals.

KEY WORDS: Biometrics · Individual recognition · Population monitoring · Conservation biology · Computational biology

—Resale or republication not permitted without written consent of the publisher—

INTRODUCTION

Accurate data on population changes and stability are vital to conservation management of threatened species (Nichols & Williams 2006). The need for identifying individual animals is essential for the understanding of population dynamics (see McMahon et al. 2007). Traditionally approaches have relied on the use of synthetic marking devices such as tarsus bands or

passive integrated transponders (PIT) but gathering long-term data sets using these techniques is labour intensive and time consuming. As a result, good demographic data are scarce (Heppell et al. 2000).

A growing number of authors and organisations recognise that the methods involved in animal research and wildlife management may not be wholly benign for the individuals or populations concerned (e.g. Cuthill 1991, Wilson & McMahon 2006).

*Email: richard.sherley@bris.ac.uk

Evidence has emerged to suggest that placing external devices on animals may influence their behaviour and fitness in natural situations (e.g. Wilson et al. 1986, Ropert-Coudert et al. 2007, Sherrill-Mix & James 2008). This may in turn influence the validity of findings from such studies; although many excellent population models exist to deal with incomplete or inadequate data sets (e.g. Hemerik & Klok 2006), the errors introduced by sampling can confound analysis (see Yoccoz et al. 2001).

The debate around the costs versus benefits of marking has been particularly vigorous when addressing the use of metal flipper bands in penguins (see Petersen et al. 2005 for review). Flipper banding has made a significant contribution to our understanding of penguin biology and has aided management since the 1950s (e.g. Sladen et al. 1968, Sidhu et al. 2007, Wolfaardt et al. 2008). However, despite advances in banding technology (Stonehouse 1999), flipper bands have been implicated in causing higher costs of swimming (Culik & Wilson 1991, Culik et al. 1993), increased foraging-trip duration (Dugger et al. 2006), lower annual return rates to breeding colonies (e.g. Clarke & Kerry 1998), substantially lower survival in the first year after banding (e.g. Froget et al. 1998) and severe flipper damage from partially opened bands (e.g. Clarke & Kerry 1998). Impacts on reproductive and fledgling success have also been shown in some species (e.g. Gauthier-Clerc et al. 2004), but not in others (e.g. Hindell et al. 1996, Barham et al. 2008), suggesting that the impact of flipper banding may vary with local environmental conditions, band design and even between species. Although the problem is complex, the balance of evidence indicates that flipper bands may not be appropriate in all circumstances and today caution is generally employed in their use (Petersen et al. 2005).

As a result, backing for the use of non-invasive methods is gaining momentum. One interesting approach is to utilise unique pattern configurations within visual markings for 'fingerprinting' individuals. Many species develop lifelong stable camouflage patterns which, in a number of cases, are believed to originate from Reaction-Diffusion systems (Turing 1952, Murray 1988). Since the formation of such systems is quasi-chaotic in nature, the resulting markings often differ widely between individuals but fit into a general theme for the species e.g. zebra stripes or cheetah spots (Murray 1988).

Traditional visual identification methods usually involve the creation of a catalogue of sketches or photographs taken by a hand-held camera (e.g. Scott 1978). The catalogue is then searched for a match by eye each time a new sighting is made (e.g. Castro & Rosa 2005). The effort for identifying large numbers of animals by these means is enormous and can take

months or years of work. Automating this process can provide a great benefit in terms of saved time and effort (Kelly 2001).

To date, only a few systems exist to aid the visual fingerprinting of animals in a wild population (e.g. Arzoumanian et al. 2005, Speed et al. 2007, Van Tienhoven et al. 2007, Gamble et al. 2008). These computer-aided approaches rely on a database of still images that must be collected manually. In instances where the species under study occurs in low abundance or is only sighted infrequently (e.g. Arzoumanian et al. 2005) attaining suitable images will generally limit the capacity of the identification system. In contrast, for colonial breeders like penguins, a fully-automated identification system could potentially populate a database of known individuals quickly and effectively. This could ultimately obviate the need for synthetic markers altogether, i.e. provide a means for autonomous, non-intrusive population monitoring.

The African penguin *Spheniscus demersus* is an ideal candidate for the application of non-invasive monitoring to the benefit of conservation. It is endemic to southern Africa, where breeding occurs between Hollamsbird Island (24° 38' S, 14° 32' E), off central Namibia, and Bird Island in Nelson Mandela Bay (33° 51' S, 26° 17' E), South Africa (Crawford & Whittington 2005). The population has decreased by at least 90 % since the turn of the 20th Century (Crawford et al. 2001) to perhaps as few as 27 000 breeding pairs today (R. J. M. Crawford unpubl. data). Good empirical data on the contemporary and on-going population fluctuations will be vital to the continued conservation of the species, particularly in light of the system wide changes currently occurring in the Benguela ecosystem (e.g. Crawford et al. 2007, 2008).

African penguins are also highly suitable for visual-identification techniques. They have high contrast plumage patterns with each individual developing unique features that remain stable for life (Reilly 1994, P. J. Barham pers. obs.). In addition, penguins move slowly and often use a limited number of predictable routes between the breeding colony and the sea, making the positioning of camera systems straightforward and giving the potential to monitor significant quantities of individuals with relatively few systems.

Burghardt and colleagues (Burghardt & Campbell 2007, Burghardt 2008) have developed and tested a prototype system capable of detecting, tracking and identifying individual African penguins based on the pattern of black spots that develops on the chest plumage. The system uses real-time video captured by a remote Ethernet camera and fed to a computer for fully-automated visual processing. Here, we provide a preliminary assessment of the efficacy of the system on

a subset of the visual data collected to date and evaluate the viability of the method as a population monitoring tool to replace more traditional approaches.

MATERIALS AND METHODS

Study site. This work was carried out at the African penguin colony on Robben Island (33° 48' S, 18° 22' E), Western Cape Province, South Africa (Fig. 1A) from April to May 2005 and April to August 2007. Robben Island was the second largest colony of African penguins in both study years, with approximately 7000 breeding pairs (Underhill et al. 2006, R. J. M. Crawford unpubl. data).

African penguins at Robben Island move to and from the sea along pathways, termed 'penguin highways'. Six major paths have been identified (T. R. Shaw unpubl. data), which correspond to 6 main landing beaches on the east coast of the island (Fig. 1B). These 6 pathway catchments serve the majority of the colony. Two of the busiest highways, termed the N1 and N2 (Fig. 1B), were selected for image acquisition. The birds at Robben Island display neophobia if novel objects are placed on their highways (R. B. Sherley pers. obs.). To reduce the influence of this behaviour, the recognition system was positioned on the highway for a number of sessions prior to the commencement of data collection until the behaviour of the penguins resembled that seen in the absence of the system.

Image acquisition. The African Penguin Recognition System (APRS) used for image capture comprises a static Ethernet camera, a Fast Ethernet infrastructure and a contemporary consumer laptop running AnimalID — a prototype, custom designed, animal recognition and identification software package (see Burghardt 2008 for technical details). The system was housed in a water- and dust-proof casing at all times whilst in the colony. The camera produces an MJPEG image stream at 6 to 8 frames per second and passes each frame (1280 × 856 pixels resolution; Fig. 2a) to the lap-top for processing.

Image acquisition took place on the N2 during 2005 and 2007 and on the N1 during 2007 only. The system was placed at a predetermined location on the selected highway just prior to the start of each session and the camera was aligned by hand to face the appropriate direction. Images were captured both in the mornings, as the birds travelled from the colony to the sea to forage, and in the afternoon, as the birds returned to their nest sites. At the end of each session, the system was recovered or repositioned as appropriate.

Image acquisition was restricted by penguin behaviour and availability of daylight (night-vision cameras were not used), thus the timings of sessions varied throughout the austral winter. Where possible, morning sessions were between 07:00 and 11:00 h from April to May and from 08:00 to 11:00 h from June to August. Afternoon sessions were from 15:30 to 19:00 h

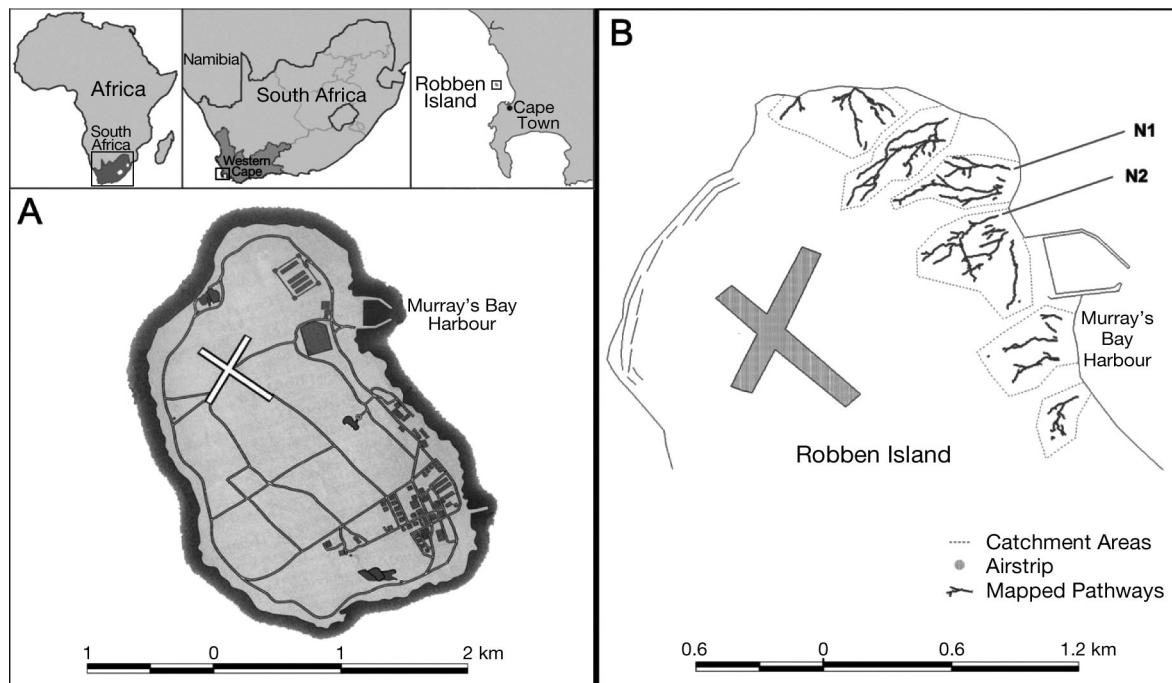


Fig. 1. (A) The location of Robben Island, within South Africa, the Western Cape Province and Table Bay. (B) The 6 main penguin highway systems of Robben Island, South Africa. The 2 penguin highways used for image acquisition, the N1 and the N2, are indicated

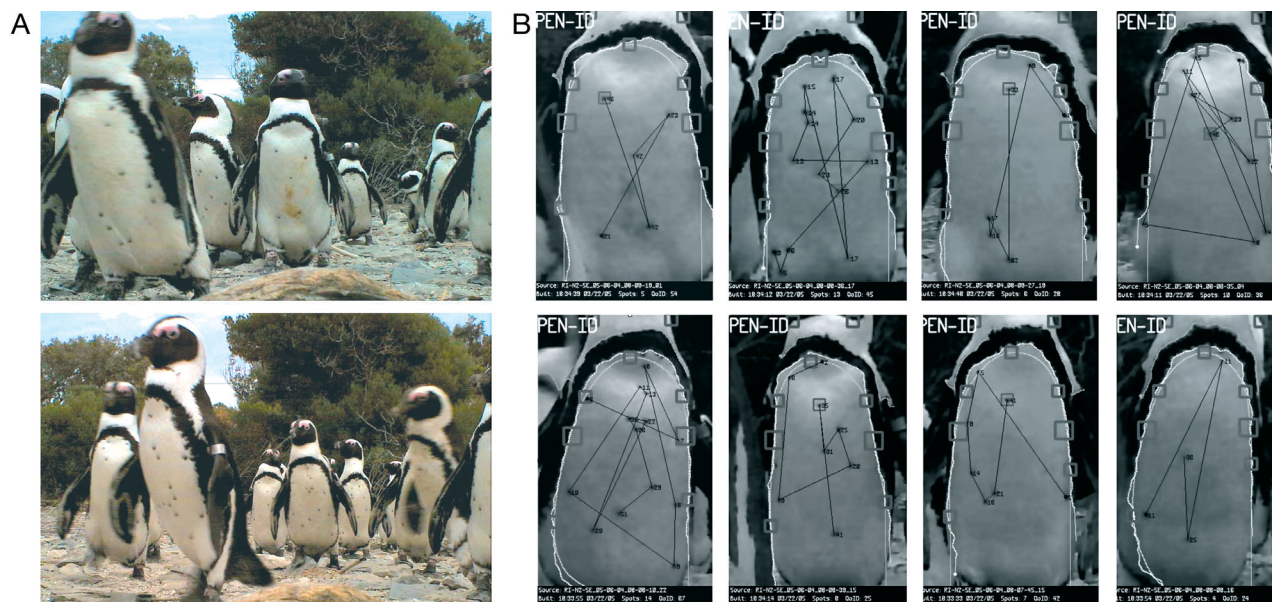


Fig. 2. (A) Representative frames of passing African penguins as captured by the field camera system; (B) sample identifications of chest patterns in the population database

during April and May and from 15:30 to 18:00 h during the rest of the year.

As each image was acquired, artificial intelligence algorithms within AnimalID would attempt to detect individual penguins in near frontal poses (where penguins align themselves so that their chests are approximately—within 20° to 30°—orthogonal to the camera axis). Sets of such frontal detections of penguins were stored together with the original image stream for subsequent further analysis (e.g. ground truthing, individual identification). The time, date, location (highway) and direction of travel was also saved for each detection stored by AnimalID.

Analysis of individual authentication and identification capacity. To estimate how reliable the software was at differentiating between individuals within a set of images, a subset of ($n = 1000$) detections of 114 birds was used to build a ground-truth data set. Each penguin detection in the subset was manually assigned to the correct individual as represented by a unique system identity. The detections used were selected at random from images acquired under favourable weather and lighting conditions (e.g. not during rain, not at dawn or dusk) during 2005.

To generate a basic benchmark, the cross-over 'authentication performance' was tested, i.e. the system's capacity to match the same individual in different images. Based on the analysis of chest patterns in AnimalID, each detection was authenticated against all other detections. The system automatically assigned either the label 'authentic' or 'different' to each of the $n(n-1)/2 = 499,500$ pairs of penguin detections in the test set.

Comparing the output of this process to the known ground truth gives the receiver operating characteristic (ROC). A key tool in pattern detection theory, a ROC comprehensively quantifies the capability of a system to perform a binary classification task: in this instance to assign the labels 'authentic' or 'different' to pairs of penguin detections. A ROC can be understood as a plot of the proportion of pairs correctly determined to be 'authentic' (the true positive or genuine acceptance rate) against the proportion of detection pairs incorrectly labelled as 'authentic' (the false positive or false acceptance rate) over the whole spectrum of possible authentication thresholds. Note that a lowered authentication threshold reduces one type of error pairs incorrectly assigned as 'different' (false negatives)—at the expense of another—pairs incorrectly labelled as 'authentic' (false positives). The optimal threshold in any one situation depends on the desired relative costs of the two sorts of error.

Based on the extraction of spot locations (see Burghardt 2008), three techniques were tested and compared for chest pattern matching. (1) Rigid alignment via Procrustes analysis (Kendall 1984) was performed to normalise globally for pattern shift, scale and rotation before calculating the mean square error (MSE) between all closest landmarks. (2) The technique of shape contexts (Belongie et al. 2000) was used to create polar histograms of spot neighbourhoods before matching the histogram sets between individuals using the Hungarian method (Kuhn 1955). (3) Finally, distortion-specific distribution contexts (Burghardt 2008) were applied in order to extend Shape Contexts by explicitly modelling the uncertainty of landmark locations governed by skin distortions.

In addition, the system's 'identification capacity' was estimated by quantifying the APRS's ability to match individuals correctly against a small sample population database. $M = 80$ of the 114 test individuals were represented as database entries (Fig. 2b), by specifying one detection of each individual as a master profile (M). The remaining 34 ind. acted as distractors (very similar non-target objects; Duncan & Humphreys 1989). The remaining 920 detections not used as master profiles were then authenticated by the system against all entries in the database, yielding either an identity via a 'match' or a 'failure to match'. Overall $M(n-M) = 73\,600$ test comparisons were performed. A comparison of the validity of the output to the ground truth provided a ROC curve for each of the three matching techniques tested. These curves describe the detailed matching capacity of the APRS for identifying individuals.

It should be noted that all the above considerations apply to the *matching* of patterns and not to the *enrolment* of new penguins. Enrolment of new birds is a separate process—it is not simply a matter of assuming that any unmatched pattern is that of a bird not already in the database.

Analysis of field monitoring capacity. To complement the data on identification performance, sample footage from the 2007 field season was analysed to determine the success and failure rates of detection. That is, whether or not an individual was detected as a species member in the first place while passing through the field of view and moving in the direction of the camera: termed making a 'camera pass'. Data from 14 sessions were selected at random, totalling ~400 000 raw images. Sessions were categorised by direction of travel ('to sea' or 'to colony'), location ('Highway N2' or 'Highway N1') and weather conditions ('sunny' or 'cloudy'). For each session the number of individual penguins that passed the camera was counted. In addition,

the 6917 detections captured by the APRS during the sessions were examined to determine the percentage of individuals that were detected amongst all camera passes. The results were summed over all sessions to yield an average likelihood that any given camera pass would result in at least one detection of the passing bird.

Finally, detections were assigned to individuals to ascertain what proportion of the individual birds passing had produced at least one detection suitable for individual recognition processing. To ensure the correct assignment of identity to all passing birds (establish ground-truth individuality), the detections were assessed by eye. A detection was deemed acceptable if both sides of the chest band were clearly visible and all spots in the pattern were visible, i.e. the bird's chest was within ~20° to 30° of laying orthogonally to the sensor and was not occluded. Consideration was also given to the lighting conditions in the detection, particularly to areas of high dynamic range. It is possible that the method produced an underestimate of the number of detections that would result in individual identification data as the AnimalID software is able to correct for some lighting variance (Fig. 3) and therefore may have processed some of the detections that were deemed unsuitable by eye. Again, the results were summed over all detections to yield an average likelihood that any given camera pass would result in at least one detection suitable for individual recognition.

To determine whether the APRS performed differently in various situations, with the aim of guiding improvements in the monitoring capacity, a binary logistic regression was used to compare the proportion of acceptable detections produced in the three different categories: direction of travel, location and weather conditions.

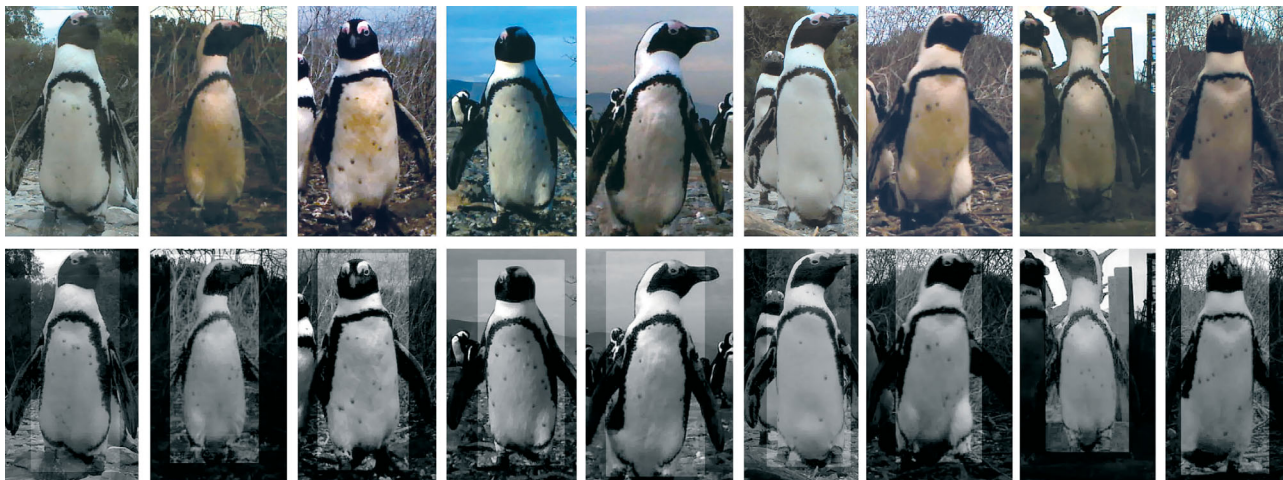


Fig. 3. Top: original penguin detection images stored by the African Penguin Recognition System (APRS); bottom: z-score lighting corrected images. Corrections can normalise for some lighting changes and should improve the ability to capture images in a field scenario

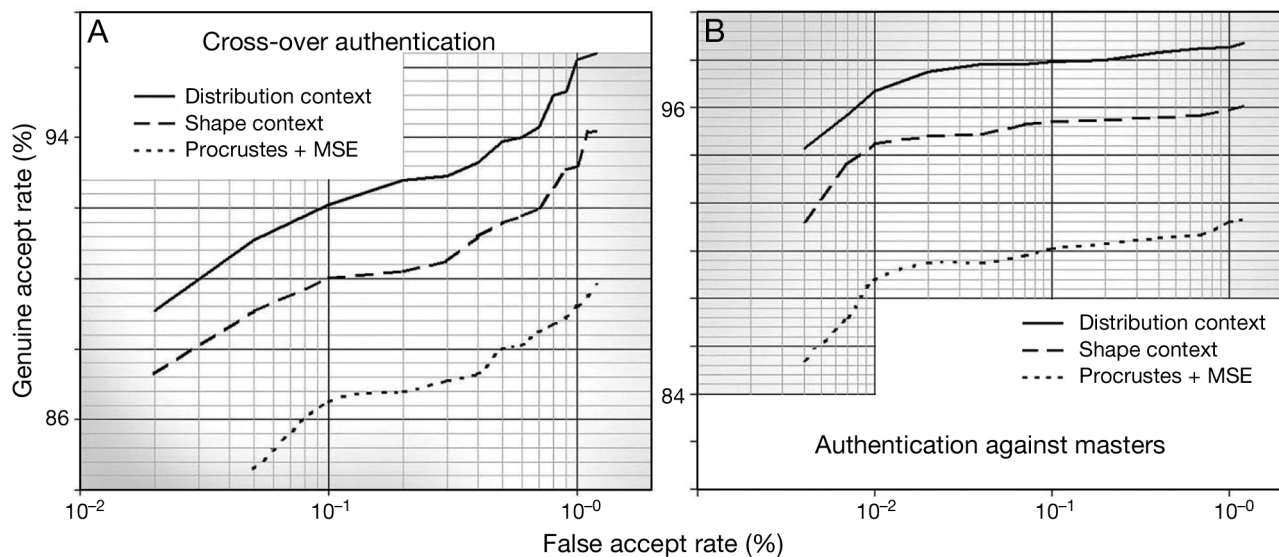


Fig. 4. African Penguin Recognition System (APRS) benchmarks for individual authentication by chest pattern. (A) Receiver operating characteristic (ROC) curves for cross-over authentication—APRS performance benchmark for matching pairs of chest image detections; (B) ROC curves for authentication against master profile—APRS performance benchmark for matching chest images against a sample population database. MSE: mean square error

RESULTS

Authentication capacity

Fig. 4A visualises the authentication performance results in the form of a ROC curve. It can be seen that a genuine acceptance rate (GAR) above 90% can be traded for a false acceptance rate (FAR) of about 3 in 10 000 (0.03%) when distribution contexts are employed. Pinpointing a practically acceptable trade-off between sensitivity and incorrect authentication, a working area on the ROC curve was stipulated at a false accept rate of 0.1%. At this point the matching software can autonomously perform at about 92% sensitivity, i.e. in 8% of all cases authentication was falsely denied.

Identification capacity

Fig. 4B shows the ROC curve for identification against a sample database containing master profiles. It clearly illustrates that a controlled master database increases accuracy and that the resulting identification system proved more robust. At an operating sensitivity of 96.7%, less than 1 in 10 000 authentications to master profiles were flawed after testing the validity of all $M_p + M_N = M$ ($n - M$) = 73 600 authentication opera-

tions. At this rate of 0.01% false acceptance, which was chosen for operation in the field, the system labelled only 7 of 72 803 negative comparisons (M_N) as positives (wrongly authenticated) whilst 26 of 797 true positives (M_p) were incorrectly rejected as matches.

Field monitoring capacity

Screening demonstrated that each time an individual penguin passed the sensor, the probability of a corresponding species detection was 0.44 with some variation seen depending on the location (Table 1). Of the 1453 birds that passed the camera during the selected sessions, 193 ind. produced detections deemed acceptable for further processing (individual identification). This represents an average probability of 0.13 for enrolling (extracting a spot pattern for matching) an individual into a population database on any given

Table 1. Results of manual screening to ascertain the likelihood that any given penguin would be recognised as a species member on any given enrolment attempt (movement past the camera). Images from 2 different locations were analysed separately; counts were then combined to produce the total estimated proportion

Location of sensor (penguin highway)	No. penguins passing sensor	No. individuals detected	Proportion detected as species member
N1	482	232	0.48
N2	971	409	0.42
Total	1453	641	0.44

camera pass, assuming that all individuals are approximately similar in their behaviour when passing the camera. To ascertain whether there were any differences that would allow for simple improvements to the system, the proportion of acceptable detections produced in the three different categories (direction of travel, location and weather conditions) were compared.

Enrolment proportion (binary logistic regression, minimally adequate model) was significantly greater on days when the weather was cloudy than on days when the weather was categorised as sunny or variable (coefficient estimate = -0.40, $z = -2.17$, $p = 0.030$). For example, the modelled enrolment probability for birds travelling towards sea on the N2 in sunny weather (0.086) was 69.4 % of the corresponding probability when conditions were cloudy (0.124). Differences were similar for the other levels of the explanatory variable (Fig. 5). Direction of travel also had a significant influence on enrolment (coefficient estimate = 1.77, $z = 6.46$, $p < 0.001$) with higher modelled probabilities when birds were travelling to sea as opposed to when they were travelling towards the colony (Fig. 5). The interaction between highway and direction of travel was also highly significant (coefficient estimate = -1.63, $z = -4.53$, $p < 0.001$; Fig. 5) probably as a result of greatly increased enrolment on the N1 when the birds were travelling to sea in both sunny and cloudy weather (Fig. 5). All other interactions and main effects failed to reach significance.

With the average likelihood of being enrolled on any given approach to the camera known, it is possible to predict the population coverage that the APRS should be able to achieve. Any individual's probability of occurring in the database can be considered as a factor of

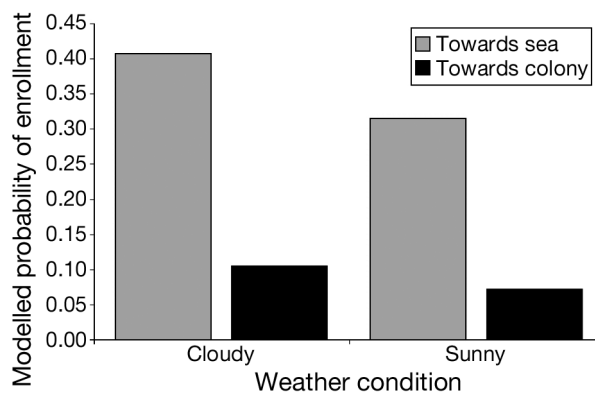


Fig. 5. Modelled probability (derived from a binary logistic regression) that a camera pass by an African penguin would result in an individual enrolment in the African Penguin Recognition System (APRS) during acquisition on the N1 highway, under two different weather conditions and both while birds walked towards the sea and towards the colony

Table 2. Predicted population coverage of the African Penguin Recognition System at the measured failure to enrol (FTE) rate of 0.87 and a theoretical FTE rate of 0.75. The coverage is dependent on the number of passes an individual makes and the FTE rate at each presentation to the camera. The coverage is shown as a proportion of birds using one highway

No. of passes	Coverage if FTE rate = 0.87	Coverage if FTE rate = 0.75
5	0.5	0.76
10	0.75	0.94
20	0.94	0.997
40	0.996	0.9999

the average failure to enrol on any occasion (1 – the average enrolment rate) and the number of times the bird passes the camera. Therefore the probability (P) of at least 1 enrolment (or resighting if the individual can be matched in the database) over any number of passes is

$$P_n[\text{enrolment}] = 1 - (P[\text{not enrolled per pass}]^n)$$

where n is the number of camera passes (Table 2).

DISCUSSION

The presented data provide a set of parameters that quantify the probability and reliability of automatic individual identifications produced by the APRS. Although the sensitivity score for the authentication sits well above the benchmarks for human voice and facial identification systems (see Ruggles 2002, Biometric Technical Assessment. www.bio-tech-inc.com/Bio_Tech_Assessment.html), penguin chest patterns lack the richness in unique information seen in some of the strongest entities exploited by human biometrics. The human iris and retina, for instance, contain thousands of features and can be used to confidently authenticate millions of individuals (Jain 2005). This level of performance will remain beyond the reaches of the APRS.

The performance outlined here represents a coarse approximation with respect to all possible application scenarios. The presented benchmarks depend on the creation of master profiles and will degenerate to some degree if the central database of individuals grows by several orders of magnitude (Jain 2005). To produce a robust performance projection for truly large-scale populations, a manual annotation of millions of samples is required.

Despite these caveats, the results from the controlled test sets provide clear evidence for the suitability of the implemented algorithms to produce an individual 'fingerprinting' of African penguins using computer-

vision techniques. The recorded matching sensitivity of 92 to 96.7% compares well with computer-aided animal biometric recognition systems currently in the literature. Such systems usually require some level of user control, either at the input level (the user must label the features to be matched by hand) or at the output level (the user is asked to select the correct match from a ranked list of images suggested by the algorithm). Where a manual choice must be made from a ranked list of between 3 and 10 potential matches, the correct individual is contained within the image subset in 84% (Ranguelova et al. 2004) to 100% of cases (e.g. Kelly 2001). However, when such systems are asked to find the 1 correct match (with no user choice), as AnimalID is asked to do, the matching accuracy drops to a low of 60.8% (Ranguelova et al. 2004) and a high of 92 to 93% (Arzoumanian et al. 2005, Van Tienhoven et al. 2007).

Obtaining high quality images that are of constant lighting and free of artefacts or shadows is, by definition, difficult in the field. Consequently, image quality is likely to be a major constraint for all visual biometric systems. For example, the inclusion of poor quality images increased false-negative matches from 6.5 to 33% in work to identify cheetahs *Acinonyx jubatus* (Kelly 2001), while Whitehead (1990) found that the probability of correctly matching two images of sperm whale *Physeter macrocephalus* flukes increased from 59 to 80% when low quality images were excluded.

The problem is made all the more acute when using automated image capture. The APRS is greatly influenced by heterogeneity in light intensity over short periods of time and at present artificial intelligence algorithms cannot respond to environmental changes as quickly or appropriately as a human photographer. During cloudy conditions the lighting was even, of low intensity and far less variable over short time spans. Shadows were also much less prominent on the birds and, consequently, probability of enrolment was much higher.

While some *post hoc* corrections can be performed on an image to normalise the effects of high dynamic range in lighting (Fig. 3), dirt patches and strong shadows will still cause the penguin detection algorithm within AnimalID to fail or will obscure spots, leading to a failure to match (Burghardt 2008). The detailed experimentation needed to physically model the specific causes of matching failure and to develop the tools needed to overcome them is outside the scope of this paper. However, this research constitutes one focal area of our current and future work. Note that advancements in camera technology may also reduce the issue of shadow-caused pattern corruption, especially with the development of high-dynamic-range cameras for the field and object-sensitive aperture

control systems. Despite these efforts, we predict that coping with the high image variability seen during environmental acquisition will remain one of the biggest challenges for automatic animal identification to overcome in the short term.

The behaviour of the animals being observed can also constrain the capacity of a monitoring system. If animals are cryptic, difficult to observe or infrequently encountered, then visual monitoring and automated image capture may be inappropriate. With the current incarnation of the APRS, individuals must appear in front of a camera, in near frontal pose and under appropriate lighting conditions. This immediately limits the population coverage to individuals using the 2 monitored highways and to those moving during daylight.

Observations in the field show that only about 40% of the birds that pass along a highway during a 24 h period do so during daylight (J. G. Underhill unpubl. data). It is not currently known whether specific individuals always restrict their movements through the colony to daylight hours or whether individuals may alter their time of arrival and departure. Whatever the case, 24 h monitoring would improve the system, either allowing for the inclusion of additional individuals in the database, or by reducing the timeframe within which individuals are resighted. Developments to modify the APRS to use infrared illumination during darkness are currently underway. This could simultaneously increase the number of birds that can be enrolled and reduce some of the problems associated with lighting, through the introduction of a controllable illumination source.

One way in which behaviour can affect the APRS was shown in the high level of potential recognition in the 'N1-towards sea' category. This seems to result from the penguins being forced to pass through a small PIT gateway just before they reached the camera. This gateway is designed to channel the birds so that they pass close to a transponder reader and it seems to have the benefit of making the birds approach the camera more directly, resulting in an increased likelihood of detection and enrolment. Without the influence of the PIT gateway, the probability of enrolment (or resighting) falls to about 0.10 on each camera pass for the other highway-direction scenarios. This suggests that it is worth considering, even where efforts are being made to monitor in a non-invasive manner, that small or benign modifications of natural behaviour may yield great benefits in terms of increased capacity when visual monitoring is employed.

Overall, the assessment of the capacity of our system in the field suggests that, even at an early stage of development, the APRS is capable of functioning as a useful tool for monitoring African penguin populations.

It has been shown that, on average, 13% of penguins can be enrolled or resighted each time they pass the camera under present operating conditions. This figure probably represents an expected lower bound for the population coverage of the system (although see discussions above about the PIT gateway) as it does not take into consideration continued monitoring over time. If the fact that birds are likely to pass the camera on more than one occasion is considered, then the likelihood of enrolling an individual over a given time frame increases.

The theoretical development of the population coverage suggests that a high percentage of birds using a monitored highway would have been captured by the APRS at least once, after they have made between 30 and 40 camera passes. If we assume (1) that all birds behave similarly as they pass the camera, (2) that they go to sea at least every other day during the breeding season and (3) that the system can be adapted so that birds using the highways during the hours of darkness can be included (see above), then a resighting frequency of about 1 observation every 1 or 2 calendar months could be expected for any individual in the population. This is probably a best case scenario and several factors can potentially limit the total population coverage. For example, with spatially fixed cameras, individuals are likely to be highly heterogeneous in their detection probability. This would limit the individual resighting frequency that the system could achieve and would need to be considered in subsequent data analysis (Borchers & Efford 2008, Royle et al. 2009).

Even with the limitations outlined above, the total population coverage that could be achieved by the APRS is likely to be significantly higher than the coverage possible with flipper banding. In 2007, only around 13% of birds at Robben Island were marked with a flipper band (R. B. Sherley unpubl. data) and, even assuming a perfect human observer and no band loss, this figure cannot be greatly increased unless further large-scale banding events take place. Using natural markings should also greatly reduce or eliminate the errors associated with band loss in historical capture-mark-recapture (CMR) studies based on banding, while the use of computer matching should reduce the errors associated with human observers, such as the misreading of band numbers. Errors in matching will occur with the APRS, particularly as the central database develops, but it should be possible to quantify these in a way that is not always possible with band reading errors. Continuous assessment of the performance parameters reported in this paper should lead to improvements in subsequent survival and population movement analyses by allowing known error rates to be fed into advanced CMR models, as has been

done for DNA-based CMR studies (e.g. Lukacs & Burnham 2005).

In addition, biometric monitoring has the advantage of being non-invasive, which may mean a benefit from reduced human presence in the colony (e.g. Ellenberg et al. 2006), the removal of any risk associated with synthetic markers and the potential bias to results associated with repeated capture of animals (see Gauthier-Clerc & Le Maho 2001). Obtaining high quality data with minimum costs to the species under study should be the ideal in animal science (Bateson 1986, Cuthill 2007) and for species in steep decline, being able to reduce the impact of monitoring may make a valuable conservation contribution in itself.

Automated biometric monitoring may well have the capacity to revolutionise terrestrial population and behavioural ecology. Visual identification techniques are widely applicable and several authors are already recognising the potential to gain data on the abundance of difficult to monitor species (e.g. Jackson et al. 2006), population dynamics (e.g. Holmberg et al. 2009) and causes of mortality (e.g. Speed et al. 2008) as well as individual behaviour (e.g. Marnewick et al. 2006). Data can be captured with something as simple as a digital camera and, as the technology improves, so the boundaries of the technique will expand. Indeed, some of the computer recognition systems currently in existence have already been adapted to studies beyond their original target species (e.g. Speed et al. 2007, Van Tienhoven et al. 2007, Burghardt 2008) and spots are not essential to the process. The algorithms underpinning APRS, for example, have already been trialled successfully on plains zebra *Equus quagga* (Burghardt 2008).

However, the method is unlikely to be a panacea. The APRS system, and animal biometrics as a whole, are still in the early days of their development and for the foreseeable future they will probably remain largely restricted to monitoring in terrestrial environments. The necessary process to confirm the multi-year stability of natural pattern on a large sample size and validate it against an unambiguous individual identifier (see Bansemer & Bennett 2008) is still ongoing for the APRS. Furthermore, any new monitoring technology should be measured against an existing, accepted technique to ensure the reliability of the identifications in a real-world scenario—as has been done with PITs in the past (e.g. Clarke & Kerry 1998). The use of animal biometrics in conjunction with other innovative assessment and monitoring methods in the ecology tool kit, such as biotelemetry and forensic techniques should lead to improvements in the data available to decision makers, allowing them to track changes in populations much more closely, more accurately and over shorter time frames. These advances

may make the difference in our challenge to reverse the global decline of many animal species by allowing decisive actions to be taken more quickly in the future.

Acknowledgements. Our research is supported by the Leverhulme Trust, the Earthwatch Institute, a National Geographic grant and our institutions. We are grateful to the Robben Island Museum for allowing us to do the fieldwork and for transportation to and from Robben Island. R. J. M. Crawford and N. J. Parsons (Marine and Coastal Management branch, Department of Environmental Affairs and Tourism, South Africa) and staff at the Animal Demography Unit (ADU), University of Cape Town provided logistical support as well as access to unpublished data. Comments from A. Steinfurth and 4 anonymous referees greatly improved an earlier version of the manuscript. L. G. Underhill provided statistical assistance and advice on presentation. T. M. Leshoro, Conservation Officer at Robben Island Museum, provided invaluable guidance and assistance in the field.

LITERATURE CITED

- Arzoumanian Z, Holmberg J, Norman B (2005) An astronomical pattern-matching algorithm for computer-aided identification of whale sharks *Rhincodon typus*. *J Appl Ecol* 42: 999–1011
- Barham PJ, Underhill LG, Crawford RJM, Leshoro TM, Bolton D (2008) Impact of flipper-banding on breeding success of African penguins *Spheniscus demersus* at Robben Island: comparisons among silicone rubber bands, stainless steel bands and no bands. *Afr J Mar Sci* 30:595–602
- Bansemmer CS, Bennett MB (2008) Multi-year validation of photographic identification of grey nurse sharks, *Carcharias taurus*, and applications for non-invasive conservation research. *Mar Freshw Res* 59:322–331
- Bateson P (1986) When to experiment on animals. *New Sci* 109:30–32
- Belongie S, Malik J, Puzicha J (2000) Shape context: a new descriptor for shape matching and object recognition. *Adv Neural Info Process Sys* 13:831–837
- Borchers DL, Efford MG (2008) Spatially explicit maximum likelihood methods for capture-recapture studies. *Biometrics* 64:377–385
- Burghardt T (2008) Visual animal biometrics: automatic detection and individual identification by coat pattern. PhD dissertation, University of Bristol
- Burghardt T, Campbell N (2007) Individual animal identification using visual biometrics on deformable coat patterns. In: The 5th International Conference on Computer Vision Systems 2007. Applied Computer Science Group, Bielefeld. <http://biocoll.uni-bielefeld.de/volltexte/2007/20>
- Castro ALF, Rosa RS (2005) Use of natural marks on population estimates of the nurse shark, *Ginglymostoma cirratum*, at Atol das Rocas Biological Reserve, Brazil. *Environ Biol Fish* 72:213–221
- Clarke J, Kerry K (1998) Implanted transponders in penguins: implantation, reliability, and long-term effects. *J Field Ornithol* 69:149–159
- Crawford RJM, Whittington PA (2005) African Penguin *Spheniscus demersus*. In: Hockey PAR, Dean WRJ, Ryan PG (eds) *Birds of Southern Africa*, 7th edn. Trustees of the John Voelcker Bird Book Fund, Cape Town, p 31–34
- Crawford RJM, David JHM, Shannon LJ, Kemper J and others (2001) African penguins as predators and prey—coping (or not) with change. *S Afr J Mar Sci* 23:435–447
- Crawford RJM, Underhill LG, Upfold L, Dyer BM (2007) An altered carrying capacity of the Benguela upwelling ecosystem for African penguins (*Spheniscus demersus*). *ICES J Mar Sci* 64:570–576
- Crawford RJM, Tree AJ, Whittington PA, Visagie J and others (2008) Recent distributional changes of seabirds in South Africa: Is climate having an impact? *Afr J Mar Sci* 30:189–193
- Culik BM, Wilson RP (1991) Swimming energetics and performance of instrumented Adélie penguins, *Pygoscelis adeliae*. *J Exp Biol* 158:355–368
- Culik BM, Wilson RP, Bannasch R (1993) Flipperbands on penguins: what is the cost of a life-long commitment? *Mar Ecol Prog Ser* 98:209–214
- Cuthill IC (1991) Field experiments in animal behaviour: methods and ethics. *Anim Behav* 42:1007–1014
- Cuthill IC (2007) Ethical regulation and animal science: why animal behaviour is not so special. *Anim Behav* 74:15–22
- Dugger KM, Ballard G, Ainley DG, Barton KJ (2006) Effects of flipper bands on foraging behaviour and survival of Adélie penguins (*Pygoscelis adeliae*). *Auk* 123:858–869
- Duncan J, Humphreys GW (1989) Visual search and stimulus similarity. *Psychol Rev* 96: 433–458
- Ellenberg U, Mattern T, Seddon PJ, Jorquera GL (2006) Physiological and reproductive consequences of human disturbance in Humboldt penguins: the need for species-specific visitor management. *Biol Conserv* 133:95–106
- Froget G, Gauthier-Clerc M, Le Maho Y, Handrich Y (1998) Is penguin banding harmless? *Polar Biol* 20:409–413
- Gamble L, Ravela S, McGarigal K (2008) Multi-scale features for identifying individuals in large biological databases: an application of pattern recognition technology to the marbled salamander *Ambystoma opacum*. *J Appl Ecol* 45:170–180
- Gauthier-Clerc M, Le Maho Y (2001) Beyond bird marking with rings. *Ardea* 89:221–230
- Gauthier-Clerc M, Gendner J-P, Ribic CA, Fraser WR, Woehler EJ, Descamps S, Gilly C, Le Bohec C, Le Maho Y (2004) Long-term effects of flipper bands on penguins. *Proc R Soc Lond B (Suppl)* 271:S423–S426
- Hemerik L, Klok C (2006) Conserving declining species using incomplete demographic information: what help can we expect from the use of matrix population models? *Anim Biol* 56:519–533
- Heppell SS, Caswell H, Crowder LB (2000) Life histories and elasticity patterns: perturbation analysis for species with minimal demographic data. *Ecology* 81:654–665
- Hindell MA, Lea MA, Hull CL (1996) The effects of flipper bands on adult survival rate and reproduction in the royal penguin, *Eudyptes schlegeli*. *Ibis* 138:557–560
- Holmberg J, Norman B, Arzoumanian Z (2009) Estimating population size, structure, and residency time for whale sharks *Rhincodon typus* through collaborative photo-identification. *Endang Species Res* 7:39–53
- Jackson RM, Roe JD, Wangchuk R, Hunter DO (2006) Estimating snow leopard population abundance using photography and capture–recapture techniques. *Wildl Soc Bull* 34:772–781
- Jain AK (2005) Biometric recognition: How do I know who you are? In: Roli F, Vitulano S (eds) *Image analysis and processing: ICIAAP 2005, lecture notes in computer science Vol 3617*. Springer-Verlag, Berlin, p 1–5
- Kelly MJ (2001) Computer-aided photograph matching in studies using individual identification: an example from Serengeti cheetahs. *J Mammal* 82:440–449

- Kendall DG (1984) Shape manifolds: procrustean metrics and complex projective shapes. *Bull Lond Math Soc* 16: 81–121
- Kuhn HW (1955) The Hungarian method for the assignment problem. *Naval Res Log Quart* 2:83–97
- Lukacs PM, Burnham KP (2005) Estimating population size from DNA-based closed capture–recapture data incorporating genotyping error. *J Wildl Manag* 69:396–403
- Marnewick KA, Bothma J du P, Verdoorn GH (2006) Using camera-trapping to investigate the use of a tree as a scent-marking post by cheetahs in the Thabazimbi district. *S Afr J Wildl Res* 36:139–145
- McMahon CR, Bradshaw CJA, Hays GC (2007) Applying the heat to research techniques for species conservation. *Conserv Biol* 21:271–273
- Murray JD (1988) How the leopard gets its spots. *Sci Am* 258: 80–87
- Nichols JD, Williams BK (2006) Monitoring for conservation. *Trends Ecol Evol* 21:668–673
- Petersen SL, Branch GM, Ainley DG, Boersma PD, Cooper J, Woehler EJ (2005) Is flipper banding of penguins a problem? *Mar Ornithol* 33:75–79
- Rangelova E, Huiskes M, Pauwels EJ (2004) Towards computer-assisted photo-identification of humpback whales. In: *Proceedings of the IEEE International Conference on Image Processing 2004*, IEEE, Piscataway, NJ, p 1727–1730
- Reilly P (1994) *Penguins of the world*. Oxford University Press, Melbourne
- Ropert-Coudert Y, Wilson RP, Yoda K, Kato A (2007) Assessing performance constraints in penguins with externally-attached devices. *Mar Ecol Prog Ser* 333:281–289
- Royle JA, Nichols JD, Karanth KU, Gopalaswamy AM (2009) A hierarchical model for estimating density in camera-trap studies. *J Appl Ecol* 46:118–127
- Scott DK (1978) Identification of individual Bewick's swans by bill patterns. In: Stonehouse B (ed) *Animal marking: recognition marking of animals in research*. MacMillan Press, London, p 160–168
- Sherrill-Mix SA, James MC (2008) Evaluating the potential tagging effects on leatherback sea turtles. *Endang Species Res* 4:187–193
- Sidhu LA, Catchpole EA, Dann P (2007) Mark-recapture-recovery modeling and age-related survival in little penguins (*Eudyptula minor*). *Auk* 124:815–827
- Sladen WJL, Wood RC, Monaghan EP (1968) The USARP bird banding program, 1958–1965. In: Austin OL (ed) *Antarctic bird studies: Antarctic research series 12*. American Geophysical Union, Washington, DC, p 213–262
- Speed CW, Meekan MG, Bradshaw CJA (2007) Spot the match—wildlife photo-identification using information theory. *Front Zool* 4:2
- Speed CW, Meekan MG, Rowat D, Pierce SJ, Marshall AD, Bradshaw CJA (2008) Scarring patterns and relative mortality rates of Indian Ocean whale sharks. *J Fish Biol* 72: 1488–1503
- Stonehouse B (1999) Penguin banding: Time for reappraisal? *Mar Ornithol* 27:115–118
- Turing AM (1952) The chemical basis of morphogenesis. *Philos Trans R Soc Lond B* 237:37–72
- Underhill LG, Crawford RJM, Wolfaardt AC, Whittington PA, and others (2006) Regionally coherent trends in colonies of African penguins *Spheniscus demersus* in the Western Cape, South Africa, 1987–2005. *Afr J Mar Sci* 28:697–704
- Van Tienhoven AM, Den Hartog JE, Reijns RA, Peddemors, VM (2007) A computer-aided program for pattern-matching of natural marks on the spotted raggedtooth shark *Carcharias taurus*. *J Appl Ecol* 44:273–280
- Whitehead H (1990) Computer assisted individual identification of sperm whale flukes. *Rep Int Whal Comm Spec Issue* 12:71–77
- Wilson RP, McMahon CR (2006) Measuring devices on wild animals: What constitutes acceptable practice? *Front Ecol Environ* 4:147–154
- Wilson RP, Grant WS, Duffy DC (1986) Recording devices on free-ranging marine animals: Does measurement affect foraging performance? *Ecology* 67:1091–1093
- Wolfaardt AC, Underhill LG, Altwegg R, Visagie J (2008) Restoration of oiled African penguins *Spheniscus demersus* a decade after the *Apollo Sea* spill. *Afr J Mar Sci* 30:421–436
- Yoccoz NG, Nichols JD, Boulinier T (2001) Monitoring of biological diversity in space and time. *Trends Ecol Evol* 16: 446–453

Editorial responsibility: Rory Wilson,
Swansea, UK

Submitted: September 10, 2009; Accepted: January 13, 2010
Proofs received from author(s): March 3, 2010