

## RESEARCH ARTICLE

# Supervised accelerometry analysis can identify prey capture by penguins at sea

 Gemma Carroll<sup>1,\*</sup>, David Slip<sup>2</sup>, Ian Jonsen<sup>2</sup> and Rob Harcourt<sup>2</sup>

## ABSTRACT

Determining where, when and how much animals eat is fundamental to understanding their ecology. We developed a technique to identify a prey capture signature for little penguins from accelerometry, in order to quantify food intake remotely. We categorised behaviour of captive penguins from HD video and matched this to time-series data from back-mounted accelerometers. We then trained a support vector machine (SVM) to classify the penguins' behaviour at 0.3 s intervals as either 'prey handling' or 'swimming'. We applied this model to accelerometer data collected from foraging wild penguins to identify prey capture events. We compared prey capture and non-prey capture dives to test the model predictions against foraging theory. The SVM had an accuracy of  $84.95 \pm 0.26\%$  (mean  $\pm$  s.e.) and a false positive rate of  $9.82 \pm 0.24\%$  when tested on unseen captive data. For wild data, we defined three independent, consecutive prey handling observations as representing true prey capture, with a false positive rate of 0.09%. Dives with prey captures had longer duration and bottom times, were deeper, had faster ascent rates, and had more 'wiggles' and 'dashes' (proxies for prey encounter used in other studies). The mean ( $\pm$ s.e.) number of prey captures per foraging trip was  $446.6 \pm 66.28$ . By recording the behaviour of captive animals on HD video and using a supervised machine learning approach, we show that accelerometry signatures can classify the behaviour of wild animals at unprecedentedly fine scales.

**KEY WORDS:** Penguin, *Eudyptula minor*, Machine learning, Support vector machine, Foraging ecology, Energetics, Feeding, Predation

## INTRODUCTION

Identifying with confidence where and when animals find food is integral to studies of foraging ecology. Accurately detecting feeding behaviour can enable energy intake to be estimated (Rothman et al., 2012), and can give insights into the distribution of food and the processes by which animals search for it (Bestley et al., 2008). However, it is difficult in practice to identify specific feeding events if directly observing the animal foraging in the wild is impractical or impossible. In these cases, biologging technologies can provide insight into the behaviour of wild animals, with the added value of simultaneously recording contextual information about the environment (Ropert-Coudert and Wilson, 2005).

Accelerometry is increasingly being used to classify behaviour states based on patterns of animal movement, and is a promising tool for identifying feeding events remotely (Lagarde et al., 2008;

Grünewälder et al., 2012; Nathan et al., 2012). Extensive high-resolution data describing continuous profiles of animal movement can now be collected reliably and cost-effectively, and the computational tools to mine this information are being accessed by ecologists (e.g. Bidder et al., 2014). Accelerometry can improve our understanding of the way that animals move through three-dimensional space and interact with the environment to acquire resources at unprecedentedly fine scales (e.g. Goldbogen et al., 2013). However, there are relatively few examples of studies in which analytical methods have been both (a) validated and (b) applied to wild animals to give insight into ecological processes (but see Nathan et al., 2012; Watanabe and Takahashi, 2013; Watanabe et al., 2014).

There are two main approaches to using accelerometry data to infer the behaviour of animals. The first is an 'unsupervised' classification approach, by which accelerometer data are grouped by similarities in movement patterns either by visual inspection of the data (in the form of a line graph) or by using techniques such as cluster analyses (Sakamoto et al., 2009) or spectral analyses (Watanabe et al., 2005; Ropert-Coudert et al., 2006). Unsupervised approaches have the benefit of being readily applicable to both new and existing datasets without the explicit need to ground-truth the information (Sakamoto et al., 2009), although validation can be done *post hoc* to confirm or improve estimates of behaviour. However, searching accelerometer data for groups of patterns related to various activities can be problematic, as this relies heavily on assumptions about how we expect animals to move and behave. For example, rapid increases in the speed of locomotion may reflect pursuit of prey, but may also signify intraspecific interaction or predator avoidance.

The second approach is 'supervised' classification, in which a model is trained on segments of movement data that have been given behaviour labels after direct observation of the animal carrying the accelerometer (Nathan et al., 2012). These ground-truthed models can then be applied to new accelerometer output to classify unobserved behaviour into pre-determined classes. Examples of this approach include machine learning techniques such as support vector machines (SVMs), classification and regression trees (CART) and artificial neural networks (ANNs), and these provide computationally powerful methods of data classification that can detect complex patterns that are not evident to the human eye. Thus, the models can identify intrinsic differences between similar behaviours or locomotory types when applied to acceleration data (Martiskainen et al., 2009). Another advantage of supervised models for determining animal behaviour is that the accuracy of the model can be tested on portions of data that are held out from model training, enabling the error rate to be clearly quantified during the model development process (Bidder et al., 2014).

Studies using accelerometry to identify feeding events in wild marine animals have favoured unsupervised models, because of the

<sup>1</sup>Department of Biological Sciences, Macquarie University, North Ryde, Sydney, NSW 2109, Australia. <sup>2</sup>Taronga Conservation Society Australia, Bradley's Head Road, Mosman, Sydney, NSW 2088, Australia.

\*Author for correspondence (gemma.carroll@mq.edu.au)

obvious logistical difficulties of ground-truthing datasets. In larger animals, accelerometers have been placed on the head, or on both the head and back to identify dynamic head movements thought to be related to 'lunging' for prey (Kokubun et al., 2011; Gallon et al., 2013; Ydesen et al., 2014) and jaw-mounted accelerometers have been used to identify mouth opening events that may result in prey ingestion (Viviant et al., 2010). In smaller marine species, back-mounted accelerometry is currently the best or only option. In these cases, periods when the wing/flipper/tail beat frequency or amplitude [identifiable from the 'heaving' (wing/flipper) or 'swaying' (tail) axis of the accelerometer data] spike above a threshold are determined to be 'dashes' – increases in speed when the animal may have encountered and pursued prey (Ropert-Coudert et al., 2006). When validation has been undertaken in the marine realm (by attaching a video camera alongside the accelerometer), it has tended to focus on validating prey capture attempts (e.g. head lunges) rather than the actual capture and consumption of prey (Watanabe and Takahashi, 2013). This distinction is important, as it is impossible to quantify an animal's foraging efficiency or its energy intake if the success of prey capture remains uncertain.

Several little penguin (*Eudyptula minor*, Forster 1781) populations around the south coast of Australia are experiencing declines that seem at least partially related to changes in the availability of the small baitfish that constitute its primary prey (Cannell et al., 2012). There is therefore an urgent need to understand the energetic requirements and feeding ecology of this species at fine scales, both to predict its vulnerability to environmental change that may alter the abundance and distribution of prey and to assess the need for conservation measures such as restrictions on fishing activity around breeding colonies. The relative ease of studying little penguins both in captivity and in the wild also makes them a suitable model species for developing analytical techniques that can be applied to other marine fauna.

Here, we developed a supervised machine learning approach to identifying feeding events at sea from observations of little penguins wearing accelerometers while they handled prey in captivity. Because of their small size, it is not yet possible to equip little penguins with cameras in the wild, and the captive setting provides a means of observing feeding behaviour in great detail. Although translating observations made in captivity to the behaviour of wild animals is inherently problematic owing to the diversity of natural behaviours and contexts, by developing a classification model based on the handling of prey rather than on pursuits or capture attempts, we can be more confident that this

approximates natural feeding behaviour. This is because the physical processes of prey handling and consumption are likely to be similar even if other predator–prey interactions may differ between captive and wild settings. By taking this approach, we also aimed to remove some of the uncertainties of prey capture variability as well as uncertainties resulting from the presence of rapid movements that are unrelated to prey encounters.

We then applied this model to wild penguin acceleration data and compared 'successful' versus 'unsuccessful' foraging dives as a means of further exploring the validity of the model on wild data, and to characterise foraging success and its determinants in little penguins. We also attempted to quantify the number of prey items ingested during a day of foraging, and validated this information in relation to what is known about little penguin energetics.

## RESULTS

### SVM performance on captive data

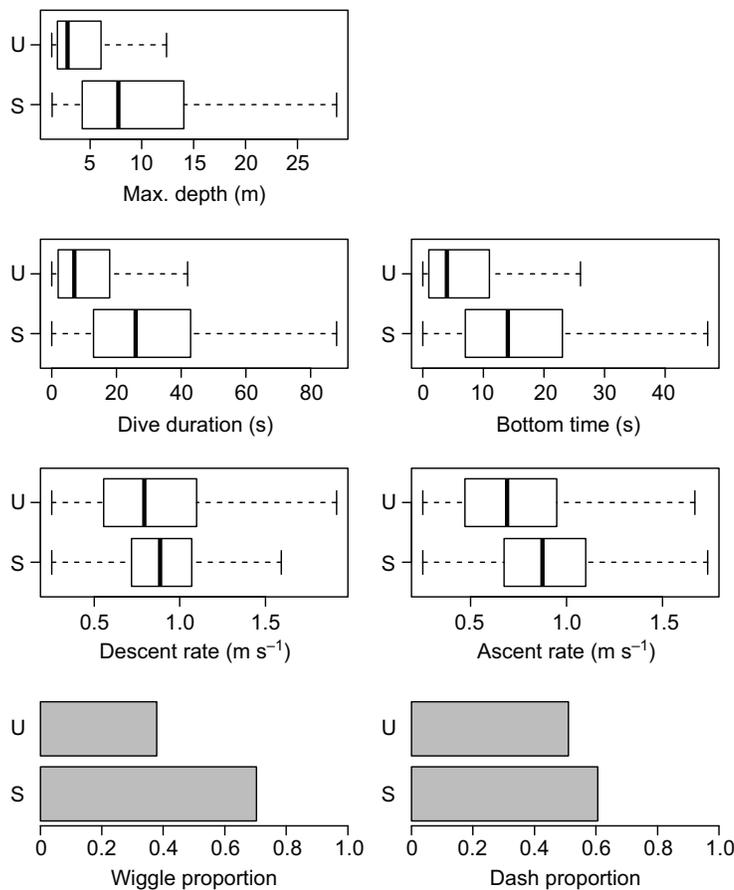
We analysed 20 accelerometry profiles recorded by seven different captive penguins on 8 days. A total of 5244 behaviour observations (3971 'swimming'; 1273 'prey handling') were extracted after processing the data into 0.3 s blocks. These observations were randomly split into 3670 training and 1574 testing points for the initial parameter tuning run, and for each subsequent iteration when evaluating the model's performance. The overall mean accuracy of the SVM in classifying both swimming and prey handling events correctly on the unseen testing data was  $84.95 \pm 0.26\%$  (mean  $\pm$  s.e.,  $N=12$  iterations). As we were interested in how well the model is likely to correctly predict prey handling in the wild, a more important measure of model performance is the false positive rate, i.e. the likelihood of misclassifying an event as prey handling when it should have been swimming. On our unseen testing data, this measure was  $9.8 \pm 0.24\%$ .

### Application of the SVM to wild accelerometer profiles

Accelerometer profiles were obtained for 21 penguins performing 23 foraging trips in 2013 (two birds were sampled twice at different times of the year and in different breeding stages – for information about penguins and their foraging trips, and measures of their foraging efficiency see Table 1). We classified a prey capture event as three consecutive observations of prey handling, in order to reduce the misclassification of transient events. As each datum is treated as an independent observation by the SVM, the probability of the model incorrectly classifying three consecutive observations as prey handling when they should all be swimming is 0.0009

**Table 1. Summary of the attributes of wild little penguins deployed with accelerometers in 2013, and measures of their foraging efficiency calculated using a support vector machine**

Month	Breeding stage	Sex	Mean mass (g)	Total dives	Prey capture dives/total dives	Prey captures/day	Prey captures/diving minute
Sep.	Incubation $N=3$ Guard $N=3$ Post-guard $N=3$ Total $N=9$	Female $N=5$ Male $N=4$	$1184 \pm 115.68$	15,557	$0.42 \pm 0.28$	$443.82 \pm 289.81$	$1.92 \pm 1.14$
Nov.	Guard $N=2$ Post-guard $N=5$ Total $N=7$	Female $N=5$ Male $N=2$	$1103 \pm 108.54$	7434	$0.37 \pm 0.23$	$465.43 \pm 454$	$1.22 \pm 0.93$
Dec.	Guard $N=6$ Post-guard $N=1$ Total $N=7$	Female $N=7$	$1043 \pm 163.78$	5797	$0.47 \pm 0.26$	$431.36 \pm 227.29$	$1.39 \pm 0.91$
Total	Incubation $N=3$ Guard $N=11$ Post-guard $N=9$ Total $N=23$	Female $N=17$ Male $N=6$	$1125 \pm 129.67$	28,788	$0.42 \pm 0.25$	$446.61 \pm 317.86$	$1.54 \pm 1.02$



**Fig. 1. A comparison of little penguin dive attributes in 2013.** Data are for dives in which the support vector machine (SVM) detected a prey capture (successful, S) and those in which it did not (unsuccessful, U).

(0.098×0.098×0.098). This is strong evidence that the prey capture events identified by the model were likely to reflect true instances when the penguin had captured and handled prey.

#### Comparison of prey capture versus non-prey capture dives

We analysed 28,788 dives and identified prey capture events in 38%. The number of prey caught per successful dive ranged from one (75% of successful dives) to six (0.02% of successful dives). We found that prey capture dives were longer in duration and bottom time, were deeper, and were more likely to contain wiggles and dashes (see Fig. 1 and below). A random subsample of 25,910 (90%) of these dives were included in the binomial generalised linear mixed model (GLMM). These results showed that dive residual (a measure of dive duration corrected for depth), mean ascent rate, wiggle presence and dash presence were all significant parameters ( $P < 0.0001$ ), while mean descent rate was not significant ( $P = 0.33$ ) (see Table 2).

**Table 2. Slope, s.e. and *P*-values for a binomial generalised linear mixed model characterising dive parameters in prey capture versus non-prey capture dives during 23 little penguin foraging trips**

Parameter	Slope	s.e.	<i>P</i>
<b>Dive residuals</b>	9.82253	0.33185	<0.0001
Mean descent rate	-0.21976	0.26772	0.334
<b>Mean ascent rate</b>	4.01958	0.15780	<0.0001
<b>Wiggle presence</b>	0.87131	0.04277	<0.0001
<b>Dash presence</b>	-0.46260	0.04259	<0.0001

Parameters in bold were significant.

#### Comparison of SVM predictions with wiggles and dashes

Wiggles (undulations in the bottom phase of a dive) occurred in 71% of prey capture dives, 40% of non-prey capture dives and 54% of all dives. Dashes (periods where the amplitude of the 'heaving' axis, a proxy for flipper strokes, spiked above a threshold) occurred in 61% of prey capture dives, 52% of non-prey capture dives and 56% of all dives. Wiggles and dashes were present together in 33% of all dives, and of these 53% were successful dives and 47% were unsuccessful dives.

#### Little penguin foraging efficiency

The number of discrete prey captures by each penguin per foraging day determined by the model was highly variable, ranging from 120 to 1368 with a mean ( $\pm$ s.e.) of  $446.61 \pm 66.28$  and a median of 305. The proportion of successful prey capture dives also varied between individuals, from 0.11 to 0.88 with a mean ( $\pm$ s.e.) of  $0.40 \pm 0.05$ . Prey captures per minute diving (averaged over the whole foraging trip) ranged from 0.30 to 4.0 with a mean ( $\pm$ s.e.) of  $1.54 \pm 0.21$  and a median of 1.56.

#### DISCUSSION

There are several emerging analytical techniques that attempt to elucidate the relationships between patterns of acceleration and behaviour in animals. Although each approach can give biological meaning to movement data, identifying and validating prey captures in a wild setting remains difficult, and very few accelerometry studies have attempted to quantify prey ingestion (Watanabe and Takahashi, 2013). The results of the present study are promising. The SVM trained on captive penguin accelerometry performed well on unseen captive data, with an 84.95% overall accuracy and 9.8%

false positive, which is similar to those found in studies on captive terrestrial animals (Nathan et al., 2012; Bidder et al., 2014). The clear differences that we identified between dives in which the SVM identified prey capture and those in which it did not are also encouraging for the translation of this technique to wild datasets. However, assessing model performance on data collected in the wild is challenging as it requires several assumptions, the validity of which we explore below.

### Prey captures and foraging strategy

Air-breathing divers such as penguins must frequently leave prey patches to return to the surface to breathe during foraging trips. This constraint confers expectations on the way that they are likely to behave while foraging. Foraging animals should make decisions that will maximise their potential for resource acquisition, and are therefore expected to spend more time foraging where food availability is high (Charnov, 1976). This prediction holds true for diving animals including wild Adélie penguins, which increase the duration of their dives in response to the rate of short-term krill capture success (Watanabe et al., 2014). Theory also predicts that shallow dives, which constituted most of the dives performed by penguins in this study (77% of all dives and 61% of successful dives were above 10 m in depth), should be aborted if prey is not encountered in the upper part of the water column (Thompson and Fedak, 2001). Therefore, successful prey capture dives are likely to be not only longer but also deeper than non-prey capture dives. The strong differences in duration and depth that we observed between prey capture and non-prey capture dives support these theories.

Penguins are observed to have sharper angles of ascent when bottom duration is longer (Sato et al., 2004), and when prey capture has been identified by oesophageal temperature loggers (Ropert-Coudert et al., 2001). This is probably because if successful prey capture dives are longer and deeper, animals are likely to be foraging closer to their aerobic dive limit and should optimise their return to the surface to replenish their oxygen stores (Viviant et al., 2014). Our results are consistent with these observations, as significantly higher mean ascent rates were observed during successful dives. Descent rate was not a determinant of prey capture in this study, indicating that there was no strong pre-conception by the penguins about whether prey was available at the single dive scale.

There was a clear relationship between the presence of both wiggles and dashes, which have been used as proxies for prey encounters in diving animals (Simeone and Wilson, 2003; Zimmer et al., 2011a), and the presence of a prey capture identified by the SVM. Wiggles were a better indicator of prey capture than dashes, corresponding to 71% of dives in which the model identified a prey capture event compared with 61% for dashes. Studies using proxies for prey encounter such as wiggles, dashes and head movements have assumed that (a) all prey that is encountered is pursued (Ropert-Coudert et al., 2006) and (b) once prey is encountered, the likelihood of prey capture is high (Zimmer et al., 2011a). However, there are many factors that are likely to affect the rate of prey capture success in relation to the prey that is encountered. These include the effects of prey patch density on prey capture success (Draulans, 1987; Darby et al., 2012), the effects of light level on the foraging success of visual predators (Ropert-Coudert et al., 2006), the presence of competition from other predators (Minderman et al., 2006) and the effects of individual experience (Daunt et al., 2007). For these reasons, rates of prey capture cannot be inferred from prey encounter, and methods that focus on prey encounters or capture

attempts alone could significantly over-estimate the prey consumed by foraging animals.

### Foraging efficiency and prey capture rates

Prey acquisition is difficult to quantify in marine animals. Techniques such as diet analyses based on stomach flushing are prone to bias as different prey types have differential digestion rates, and hard parts such as otoliths and squid beaks, if consumed, may not always be retained in the gut (Gales, 1988). For animals performing longer foraging trips, stomach contents on return to the colony may only represent the most recent bout of feeding. Such studies generally report the relative occurrence of prey types, rather than attempting to quantify the number of prey items that have been ingested. However, approximate quantities can be estimated by back-calculating the wet mass of key prey types required to fulfil the daily energy requirements of individuals based on known metabolic rates (Adams et al., 1993; Watanabe and Takahashi, 2013) or they can be determined via water and sodium influxes (Gales and Green, 1990).

To determine whether the SVM result provides a reasonable estimate of the quantity of prey items ingested, we estimate the possible feeding requirements of little penguins from what is known about their energetics. The SVM identified a mean of 447 prey capture events per penguin per day. The relationship between the daily metabolic rate of little penguins during the breeding season and the food that they require is complicated by the penguins' need to consume a surplus to account for fasting days and the extra mass of food required to provision growing chicks. The metabolic rate of little penguins during the breeding season has been determined to be between  $661 \text{ kJ kg}^{-1} \text{ day}^{-1}$  (for a fasting, incubating bird) and  $2532 \text{ kJ kg}^{-1} \text{ day}^{-1}$  (for a foraging bird late in the chick-provisioning phase) using the doubly labelled water method (Gales and Green, 1990). The mean energy value of the little penguin diet, with varying proportions of fish, squid and krill across the annual cycle is  $\sim 3.87 \text{ kJ g}^{-1}$ . Food consumption based on water and sodium influx has been calculated as being up to  $664 \text{ g kg}^{-1} \text{ day}^{-1}$  (Gales and Green, 1990).

Multiplying the standardised metabolic rate by the mean mass of the penguins in this study (1.125 kg), foraging penguins provisioning chicks require up to  $2848 \text{ kJ day}^{-1}$ . This approximate energy requirement could be met with 736 g of food, a similar mass to the 747 g obtained when the mass of food estimated by water and sodium influx is also multiplied by 1.125 kg (Gales and Green, 1990). Food consumption has also been estimated for little penguin chicks using water and sodium influx, with an estimated mean rate of intake of  $256 \text{ g day}^{-1}$  for older (post-guard stage) chicks (Green et al., 1988). This means that  $\sim 1259 \text{ g}$  of food could be eaten on an average day by a foraging penguin provisioning two chicks. This estimate may be lower for incubating birds and those with very small chicks; however, this figure provides a benchmark on which to sanity test our observations.

If all prey captures identified by the SVM represent actual prey captures, prey consumed by little penguins in our study had a mean mass of  $\sim 2.8 \text{ g}$ . The mean mass of ingested fish species estimated from diet reconstruction for little penguins at Phillip Island, Victoria, Australia, ranged from 0.9 g (red cod) to 10.7 g (garfish), with a mean of  $\sim 4.9 \text{ g}$  per item across all fish species (Cullen et al., 1991). Small (1–5 cm) schooling fish were the primary component of the diet of little penguins around Montague Island in 2013, with penguins also eating krill when transient swarms entered the area (G.C., unpublished data). Our estimate of mean prey ingestion quantities is consistent with this provided that the handling of small

size classes of fish and krill is considered in the same way as the handling of larger fish by the SVM.

As the model was trained on penguins handling one type of prey (pilchards), whether or not the capture of other prey types can be identified with the same accuracy is uncertain. This is a limitation of the interpretability of non-linear SVMs, which are to some extent ‘black box’ algorithms with non-transparent decision rules (Rosenbaum et al., 2011). Logically, the capture and handling of any prey type (live or dead) is more likely to resemble the physical action of handling dead pilchards than of swimming. It is probable that when the accelerometer records these events, the SVM classifies the observations as ‘prey handling’, and the number of prey captures recorded in our study supports the notion that the capture of small prey items is being detected.

However, in this study we made a number of assumptions regarding the amount of time it takes for a little penguin to handle prey in the wild. By setting a prey capture event at three consecutive 0.3 s observations of prey handling, the minimum observable prey handling time was 0.9 s, which may not be at a fine enough resolution to detect the rapid consumption of small prey. Similarly, by using a survival curve to determine the amount of time between consecutive prey captures it is possible that we also underestimated prey capture rates when prey density was high. Without validating the handling of more prey types (particularly small fish and krill) and of both dead and live prey, there will remain a level of uncertainty, and the absolute estimates of prey capture quantity should be interpreted with caution at this early stage.

### Benefits of supervised accelerometry analysis and future applications

The ability to determine the quantity and timing of prey consumption in wild animals has immense value in ecological studies. The supervised accelerometry approach provides a reliable and minimally invasive method to assess and monitor foraging efficiency, and to estimate the energy intake of predators. Further value can be added by integrating this information with fine-scale location data and environmental covariates such as temperature and light level (Guinet et al., 2014). This will lead to an improved understanding of the factors governing the distribution of resources in both space and time, and enable better predictions of the effects of environmental change on predator populations (Hazen et al., 2013).

Although identifying and quantifying the consumption of prey gives us valuable information in the little penguin example, supervised accelerometry analysis can be equally applied to other behaviours of interest. We have shown here that fine-scale behaviours that correspond to transient events lasting less than a second can be detected accurately with a machine learning model such as an SVM. Therefore, this technique is potentially valuable for identifying other behaviours of interest that are variable or transient, or occur relatively rarely, e.g. aggressive interactions or copulation.

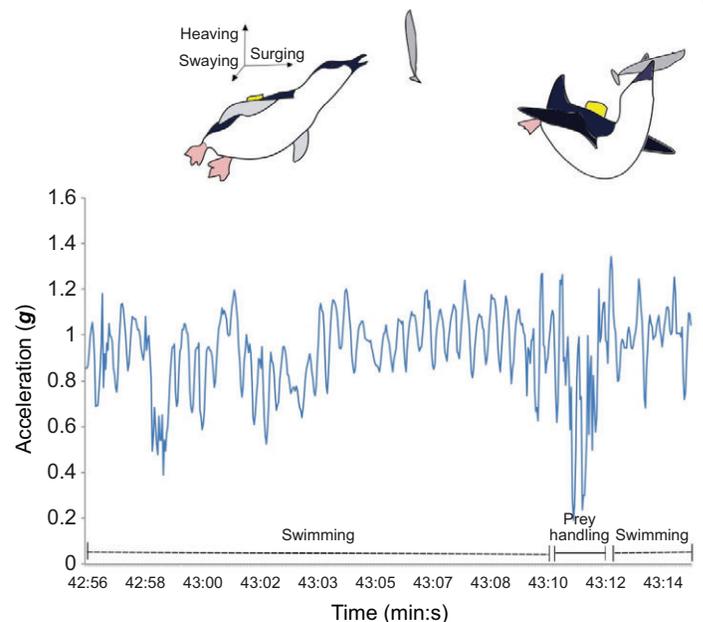
In order to detect events at such a fine scale, the accelerometry data must be high resolution (30 Hz or greater) (Broell et al., 2013) and behaviour analysis must occur at the same resolution in order to pinpoint the transition between behaviour states with precision. Ground-truthing accelerometry data in captivity is valuable in this respect, as high-quality video can be collected from angles that clearly show the whole organism and all its movements to allow for more precise coding of behaviour. The opportunity also exists to experimentally increase the frequency of a particular activity, whereas data must be collected in a much more opportunistic

fashion in the wild. Although critter cams are valuable for gaining insight into the natural behaviours of wild animals, they may not always show a behaviour from the best angle, making it difficult, for example, to determine whether prey capture was successful (Watanabe and Takahashi, 2013). They are also a sub-optimal method for animals that operate primarily in low light conditions such as nocturnal or deep diving animals, because of the potential interference of the camera’s light source with the animal’s normal behaviour (Heaslip and Hooker, 2008). Ideally, observation of animals in the wild is the best way to remove doubts regarding the validity of acceleration signatures for identifying the full range of their natural behaviours. However, if this is impractical and if captive animals perform behaviours that adequately reflect the natural behaviour of their wild conspecifics, the captive environment provides a useful setting to begin to model their fine scale behaviour using accelerometry.

### MATERIALS AND METHODS

The field study was conducted on Montague Island (−36.252777°, 150.227110°), 9 km off the southeast coast of NSW, Australia. The island supported approximately 5000 breeding pairs of little penguins in 2000 (Weerheim et al., 2003); however, there have been no recent, reliable estimates of population size on the island. Extensive habitat restoration on Montague Island has necessitated the use of nest boxes to accommodate a large portion of the breeding penguin population. Data used in this analysis were collected during the breeding season in September, November and December 2013. Penguins were captured in their artificial wooden nest boxes at night, and sex was determined by comparing the morphology of the two adults present in the nest (males are generally slightly larger in body size and have thicker bills with a pronounced hook at the tip).

Accelerometer data loggers (G6a and G6a+, CEFAS Technology Pty Ltd, Suffolk, UK) were attached to feathers just below the mid-point of the back with cloth tape (Tesa, Hamburg, Germany) the night before a penguin went to sea. The units recorded acceleration in three axes: anterior–posterior (surging), lateral (swaying) and dorso-ventral (heaving) with a range of  $\pm 2$  g (see Fig. 2). The accelerometers recorded depth, temperature and



**Fig. 2. Schematic diagram of a little penguin wearing an accelerometer (yellow) swimming towards a fish and handling prey.** The accelerometer axes (heaving, surging and swaying; see Materials and methods) are shown on the left. Beneath is a sample raw accelerometry profile from the heaving axis (recorded at 30 Hz) of a penguin swimming and handling prey in captivity, labelled with the associated behaviours identified from HD video.

acceleration and were programmed in two modes: 'shallow' mode (<1.37 m: 1.5% of the full scale pressure range), where parameters were recorded every 10 s; and 'dive' mode (>1.37 m), where the same parameters were recorded at a rate of 30 Hz. When the penguins returned from a foraging trip, they were recaptured in their nest boxes, the loggers were removed and the penguins were weighed in a calico bag using a spring balance scale (Pesola AG, Switzerland).

### Validation of accelerometer signal

To identify a unique signal from acceleration data that corresponds to prey handling in little penguins, we attached the same accelerometers to captive penguins at Taronga Zoo, Sydney, Australia, using the same technique as for wild penguins. The accelerometers were programmed to record depth and tri-axial movement continuously at 30 Hz. Three underwater cameras (GoPro Hero 3, San Mateo, CA, USA) filming in HD 1080 at 60 frames s<sup>-1</sup> were fixed in the pool and angled to give coverage of the entire swimming area. The accelerometers were attached using uniquely coloured cloth tape so that each penguin with an accelerometer could be individually identified from the video. First, we attached accelerometers to five penguins and recorded their behaviour as they swam around their enclosure with other members of their captive group for 1 h. In subsequent experiments, accelerometers were attached to two to five penguins in the morning, and were programmed to start recording before the afternoon feeding session. Two GoPro cameras were then strategically placed to provide coverage of the feeding area at the start of each feeding session.

The behaviour of the penguins was determined from the video footage and recorded directly onto the accelerometer output file. This created a behaviour label for each 30 Hz accelerometer reading. Exact pairing of the video footage with accelerometer data to the 30 Hz level was achieved by a combination of (a) identifying the exact frame within a given second in which a change in behaviour occurred on the video and (b) visually identifying rapid changes in acceleration associated with a change in behaviour from plotted accelerometer data. Behaviour was scored as being one of 'swimming', 'surface swimming' or 'prey handling'.

Because of ethical and practical considerations, live prey was not used in feeding trials. Instead, handfuls of dead pilchards were thrown into the pool until the penguins were satiated. Prey handling was recorded from the time that the penguin grabbed a fish in the water until the first powerful flipper stroke as it began to swim away after swallowing the fish. We believe that this reasonably approximates the prey handling behaviour of wild penguins, as: (1) the captive penguins approached a fish at high speed and with dynamic movement as there were usually several penguins competing for each fish and (2) the captive penguins performed characteristic behaviours such as striking at the head of the fish to 'immobilise' it before swallowing it head first. All other behaviour, including all feeding behaviour leading right up to grabbing a fish, was recorded as swimming unless the penguin was at the surface. All observations where the penguin was recorded as surface swimming were later removed from the analysis, as the wild data did not capture times when the birds rested at the surface at the same 30 Hz resolution.

### Data analysis

#### SVM and data processing

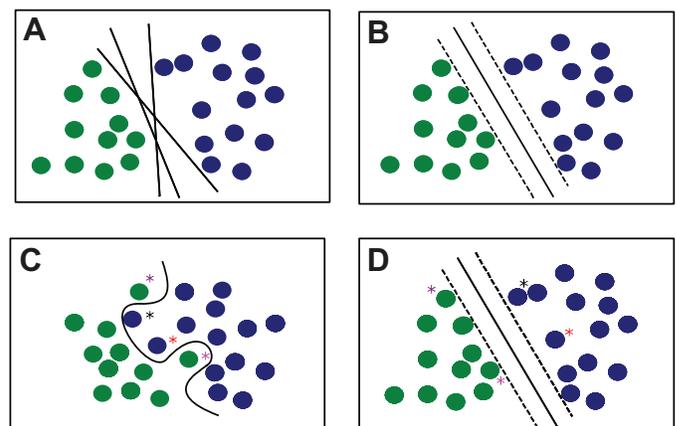
Twenty-eight summary statistics were calculated from the raw accelerometer output using a rolling window of 10 data points (0.3 s). For each axis (heaving, surging and swaying), we calculated mean, s.d., minimum, maximum, skewness and kurtosis. We also calculated pairwise correlations between the three axes as well as overall dynamic body acceleration, an estimate of activity-specific metabolic rate (Wilson et al., 2006). Rather than pre-segmenting the data into groups representing single behaviours for the training/testing process (e.g. Nathan et al., 2012), we took the behaviour label of each 10 data point rolling window to be whichever behaviour was represented by most (>5) observations. Having behaviour inputs that are mixed in this way introduces uncertainty and is therefore likely to reduce the accuracy of our model when tested on captive data. However, this should result in a model that can be applied with increased confidence to wild accelerometry, where the model must be robust to data that are not grouped into classes.

Several powerful supervised machine learning methods have been successfully employed to classify the behaviour of terrestrial animals from accelerometry, including classification and regression trees, random forests, artificial neural networks and SVMs. Each of these models performs to a high standard when classifying animal behaviour (see Nathan et al., 2012; Martiskainen et al., 2009). We selected the SVM to classify feeding because it is robust and consistently among the best performers in comparisons of machine learning methods. Rather than comparing a variety of statistical methods, we focus instead on exploring the ability of a single method to give insight into fine scale ecological processes.

SVMs are used in many pattern recognition applications and we have attempted to describe the underlying mechanisms of the model here in relation to classifying animal behaviour [for a detailed explanation, see Shawe-Taylor and Christianini (Shawe-Taylor and Christianini, 2004)]. As in all supervised machine learning problems, the model is trained on data that have been given class labels. In this case, the data are the summary statistics derived from the acceleration data, and the labels are the behaviours that the animal was observed performing from the video that correspond to the acceleration data. The model 'learns' the relationship between features of the data and these behaviour labels, and is then able to classify new unlabelled data into these classes.

The SVM can be visualised conceptually in two dimensions, with points representing two linearly separable classes, e.g. swimming and prey handling (see Fig. 3). There are many lines that could be drawn that would separate the data into these two classes. The SVM algorithm aims to determine the line (or 'hyperplane') that is able to separate the data with the largest possible margin. This large margin makes the model robust to new data that may fall outside the range of the observations used to train the model. The optimal hyperplane determined during the training process is then used to classify new examples.

Rather than classifying data in two dimensions, in reality the space is *n*-dimensional, where each dimension represents a different user-specified feature of the data. In this case, features could be any number of summary statistics derived from the raw accelerometer data across some time window (e.g. the mean of the heaving accelerometer axis over 0.3 s) and many features are usually required for the model to accurately detect patterns. Unsurprisingly, separating the two classes with a linear classifier is often not possible. A kernel represented by  $k(x_i, y_j)$ , where  $x_i$  is the sample vector input as training data, and  $y_j$  is the class label of  $x_i$ , is therefore employed to increase the distance between the classes by transforming or 'mapping' them in a high dimensional feature space. The type of kernel that is selected defines this feature space.



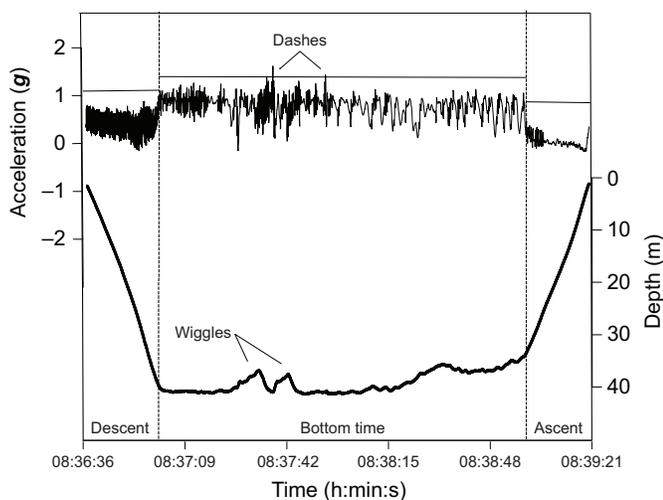
**Fig. 3. Conceptual 2D representation of how SVMs separate data into two classes.** (A) A series of possible lines that could separate data of two classes; (B) the optimal line or 'hyperplane' determined by a support vector machine that separates the data with the largest margin; (C) data that cannot be linearly separated; and (D) the effect of applying a 'kernel trick' to increase the distance between the classes in C by projecting them in hyperspace. Points with the same coloured asterisk in C and D represent the same data point.

We trained an SVM to classify the behaviour of the penguins as swimming or prey handling from the labelled data in the freely available R statistical software (R Development Core Team, 2013) package e1071 (Meyer et al., 2014). The data were randomly split into two sets in a 70:30 ratio (training:testing). Tuning of the SVM parameters was performed using 10-fold cross-validation. We trialled a range of available kernels (radial, linear, second-, third- and fourth-order polynomial kernels), and selected a second-order polynomial kernel [represented by  $k(x_i, y_i) = (x_i, y_i)^2$ ] as the model developed with this kernel showed the highest overall accuracy. Mean and s.d. of the overall accuracy and false positive rate, a measure of how often the model misclassifies a point as prey handling when it should be swimming (incorrect 'prey handling'/incorrect 'prey handling' + true 'swimming'), of the best model were obtained by generating random 70:30 splits, re-training and testing the accuracy of the model on each of these datasets.

#### Application of SVM to wild accelerometer dataset

When the best model had been selected, we then applied it to the wild dataset, after pre-processing the data using the same statistics and 10 data point (0.3 s) rolling window. We classified a 'prey capture event' as any sequence of three or more consecutive prey handling classifications, to reduce the misclassification of transient events. We determined a threshold for the amount of time that separated discrete feeding events by plotting a survival curve showing the frequency of seconds (1–100 s) between prey handling events for each individual penguin. The breakpoint of the curve was used to define the threshold, which was between 5 and 9 s for all penguins. We took the timestamp of these feeding events (to the nearest second) to be the 5th data point of the sequence of 10 data points used in calculating the summary statistics.

We calculated a number of summary statistics in order to characterise intrinsic differences between dives in which the SVM identified prey capture and those in which it did not. We first calculated maximum depth, dive duration, bottom time duration, mean ascent and descent rates, wiggle presence and dash presence. We included wiggles and dashes in this part of the analysis as they have been used as proxies for prey encounter in previous studies, and we wished to understand whether their presence was related to predictions of prey capture derived from the SVM. Dashes are instances where the penguin increases the amplitude of its flipper strokes, possibly to pursue prey (Ropert-Coudert et al., 2006; Zimmer et al., 2011a). For the dash analysis, we identified an upper amplitude threshold for flipper strokes (identifiable in the vertical heaving axis) by plotting a survival curve of amplitude frequency from the raw accelerometer data. We calculated



**Fig. 4. Depth and acceleration in the vertical heaving axis during a sample little penguin dive that included both wiggles and dashes.** Wiggles are undulations in the bottom phase of the dive occurring at  $>0.5 \text{ m s}^{-1}$ ; dashes are spikes above an acceleration threshold determined using a survival curve. Both wiggles and dashes have been used as proxies for prey encounter and are included in this paper to understand their relationship with predictions from the SVM estimates of prey capture.

thresholds separately for the descent, bottom time and ascent phases of dives and for each individual bird. As penguins must stroke harder in the top part of the water column to compensate for buoyancy, the descent phase of a dive was not analysed if the depth was less than 4 m (Zimmer et al., 2011b).

Wiggles are undulations in the bottom phase of dives thought to be related to hunting strategy, and have been shown to be a proxy for prey encounter in other species of penguins (Simeone and Wilson, 2003; Bost et al., 2007) and whales (Goldbogen et al., 2013). We defined a wiggle as a change in depth during bottom time occurring at  $>0.5 \text{ m s}^{-1}$  (see Fig. 4). We defined the start and end of bottom time as the first and last time within a dive that the rate of change in depth was  $<0.25 \text{ m s}^{-1}$  (Kato et al., 2008).

To determine which of these features are important components of prey capture versus non-prey capture dives, we used a binomial GLMM in the R package lme4 (Bates et al., 2014). As dive duration, maximum depth and bottom duration were all strongly, linearly correlated, we calculated the 'dive residual': a measure of dive duration after removing the effect of maximum depth, and dropped both maximum depth and dive duration from the model. The dive residual is likely to be a proxy for bottom duration, with removal of the contribution of time spent travelling to and from depth. We therefore also omitted bottom time duration from the GLMM. Fixed effects were dive residual, mean ascent rate, mean descent rate, wiggle presence and dash presence. As these variables are naturally autocorrelated at the single dive scale (the presence or absence of prey capture in one dive is not an independent observation if the penguin is foraging in bouts), we ran the GLMM on a random sub-sample representing 90% of the dives, to reduce the autocorrelation. Penguin ID was included as a random effect.

#### Acknowledgements

We would like to acknowledge the support of Taronga staff, including Ryan Tate, Jose Altuna and Kat Theodore for their time and assistance in the captive feeding experiments. We would also like to thank the staff at NSW National Parks and Wildlife Service at Narooma for administrative and field support, and Dr Ben Pitcher for his assistance in the 2013 field season.

#### Competing interests

The authors declare no competing financial interests.

#### Author contributions

G.C., D.S. and R.H. designed and carried out the zoo experiments and field study. G.C. designed and conducted the analyses with statistical advice from I.J. G.C. prepared the manuscript with input from D.S., R.H. and I.J. on the interpretation and presentation of results.

#### Funding

This project is funded by Australian Research Council Linkage Grant [grant number LP110200603], with support from Taronga Conservation Society Australia. G.C. is a recipient of a Macquarie University Research Excellence Scholarship. I.J. is funded by a Macquarie Vice-Chancellor's Innovation Fellowship.

#### References

- Adams, N. J., Moloney, C. and Navarro, R. (1993). Estimated food consumption by penguins at the Prince Edward Islands. *Antarct. Sci.* **5**, 245–252.
- Bates, D., Maechler, M., Bolker, B., Walker, S., Haubo Bojesen, R. Christensen, Singmann, H. and Dai, B. (2014). *lme4: Linear, Mixed-effects models using Eigen and S4*. R package version 1.1-6. Available at: <http://CRAN.R-project.org/package=lme4>
- Bestley, S., Patterson, T. A., Hindell, M. A. and Gunn, J. S. (2008). Feeding ecology of wild migratory tunas revealed by archival tag records of visceral warming. *J. Anim. Ecol.* **77**, 1223–1233.
- Bidder, O. R., Campbell, H. A., Gómez-Laich, A., Urgé, P., Walker, J., Cai, Y., Gao, L., Quintana, F. and Wilson, R. P. (2014). Love thy neighbour: automatic animal behavioural classification of acceleration data using the k-nearest neighbour algorithm. *PLoS ONE* **9**, e88609.
- Bost, C. A., Handrich, Y., Butler, P. J., Fahlman, A., Halsey, L. G., Woakes, A. J. and Ropert-Coudert, Y. (2007). Changes in dive profiles as an indicator of feeding success in king and Adélie penguins. *Deep Sea Res. Part II Top. Stud. Oceanogr.* **54**, 248–255.
- Broell, F., Noda, T., Wright, S., Domenici, P., Steffensen, J. F., Auclair, J. P. and Taggart, C. T. (2013). Accelerometer tags: detecting and identifying activities in fish and the effect of sampling frequency. *J. Exp. Biol.* **216**, 1255–1264.
- Cannell, B. L., Chambers, L. E., Wooller, R. D. and Bradley, J. S. (2012). Poorer breeding by little penguins near Perth, Western Australia is correlated with above average sea surface temperatures and a stronger Leeuwin Current. *Mar. Freshw. Res.* **63**, 914–925.

- Charnov, E. L. (1976). Optimal foraging, the marginal value theorem. *Theor. Popul. Biol.* **9**, 129-136.
- Cullen, J. M., Montague, T. L. and Hull, C. (1991). Food of little penguins *Eudyptula minor* in Victoria: comparison of three localities between 1985 and 1988. *Emu* **91**, 318-341.
- Darby, P. C., Fujisaki, I. and Mellow, D. J. (2012). The effects of prey density on capture times and foraging success of course-hunting adult snail kites. *Condor* **114**, 755-763.
- Daunt, F., Wanless, S., Harris, M. P., Money, L. and Monaghan, P. (2007). Older and wiser: improvements in breeding success are linked to better foraging performance in European shags. *Funct. Ecol.* **21**, 561-567.
- Draulans, D. (1987). The effect of prey density on foraging behaviour and success of adult and first-year grey herons (*Ardea cinerea*). *J. Anim. Ecol.* **56**, 479-493.
- Gales, R. and Green, B. (1990). The annual energetics cycle of little penguins (*Eudyptula minor*). *Ecology* **71**, 2297-2312.
- Gales, R. P. (1988). The use of otoliths as indicators of Little Penguin *Eudyptula minor* diet. *Ibis* **130**, 418-426.
- Gallon, S., Bailleul, F., Charrassin, J. B., Guinet, C., Bost, C. A., Handrich, Y. and Hindell, M. (2013). Identifying foraging events in deep diving southern elephant seals, *Mirounga leonina*, using acceleration data loggers. *Deep Sea Res. Part II Top. Stud. Oceanogr.* **88-89**, 14-22.
- Goldbogen, J. A., Calambokidis, J., Friedlaender, A. S., Francis, J., DeRuiter, S. L., Stimpert, A. K., Falcone, E. and Southall, B. L. (2013). Underwater acrobatics by the world's largest predator: 360° rolling manoeuvres by lunge-feeding blue whales. *Biol. Lett.* **9**, 20120986.
- Green, B., Brothers, N. and Gales, R. (1988). Water, sodium and energy turnover in free-living little penguins, *Eudyptula minor*. *Aust. J. Zool.* **36**, 429-440.
- Grünewälder, S., Broekhuis, F., Macdonald, D. W., Wilson, A. M., McNutt, J. W., Shawe-Taylor, J. and Hailes, S. (2012). Movement activity based classification of animal behaviour with an application to data from cheetah (*Acinonyx jubatus*). *PLoS ONE* **7**, e49120.
- Guinet, C., Vacuqué-García, J., Picard, B., Bessigneul, G., Lebras, Y., Dragon, A. C., Viviant, M., Arnould, J. P. Y. and Bailleul, F. (2014). Southern elephant seal foraging success in relation to temperature and light conditions: insight into prey distribution. *Mar. Ecol. Prog. Ser.* **499**, 285-301.
- Hazen, E. L., Jorgensen, S., Rykaczewski, R. R., Bograd, S. J., Foley, D. G., Jonsen, I. D., Shaffer, S., Dunne, J. P., Costa, D., Crowder, L. B. et al. (2013). Predicted habitat shifts of Pacific top predators in a changing climate. *Nat. Clim. Change* **3**, 234-238.
- Heaslip, S. G. and Hooker, S. K. (2008). Effect of animal-borne camera and flash on the diving behaviour of the female Antarctic fur seal (*Arctocephalus gazella*). *Deep Sea Res. Part I Oceanogr. Res. Pap.* **55**, 1179-1192.
- Kato, A., Ropert-Coudert, Y. and Chiaradia, A. (2008). Regulation of trip duration by an inshore forager, the little penguin (*Eudyptula minor*), during incubation. *Auk* **125**, 588-593.
- Kokubun, N., Kim, J. H., Shin, H. C., Naito, Y. and Takahashi, A. (2011). Penguin head movement detected using small accelerometers: a proxy of prey encounter rate. *J. Exp. Biol.* **214**, 3760-3767.
- Lagarde, F., Guillon, M., Dubroca, L., Bonnet, X., Ben Kaddour, K. and Slimani, T. (2008). Slowness and acceleration: a new method to quantify the activity budget of chelonians. *Anim. Behav.* **75**, 319-329.
- Martiskainen, P., Järvinen, M., Skön, J. P., Tiirikainen, J., Kolehmainen, M. and Mononen, J. (2009). Cow behaviour pattern recognition using a three-dimensional accelerometer and support vector machines. *Appl. Anim. Behav. Sci.* **119**, 32-38.
- Meyer, D., Dimitriadou, E., Hornik, K., Weingessel, A. and Leisch, F. (2014). *e1071: Miscellaneous Functions of the Department of Statistics (e1071)*, TU Wien. R package version 1.6-2. <http://CRAN.R-project.org/package=e1071>
- Minderman, J., Lind, J. and Cresswell, W. (2006). Behaviourally mediated indirect effects: interference competition increases predation mortality in foraging redshanks. *J. Anim. Ecol.* **75**, 713-723.
- Nathan, R., Spiegel, O., Fortmann-Roe, S., Harel, R., Wikelski, M. and Getz, W. M. (2012). Using tri-axial acceleration data to identify behavioral modes of free-ranging animals: general concepts and tools illustrated for griffon vultures. *J. Exp. Biol.* **215**, 986-996.
- R Development Core Team (2013). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. Available at: <http://www.R-project.org/>.
- Ropert-Coudert, Y. and Wilson, R. P. (2005). Trends and perspectives in animal-attached remote sensing. *Front. Ecol. Environ.* **3**, 437-444.
- Ropert-Coudert, Y., Kato, A., Baudat, J., Bost, C. A., Le Maho, Y. and Naito, Y. (2001). Time/depth usage of Adélie penguins: an approach based on dive angles. *Polar Biol.* **24**, 467-470.
- Ropert-Coudert, Y., Kato, A., Wilson, R. P. and Cannell, B. (2006). Foraging strategies and prey encounter rate of free-ranging little penguins. *Mar. Biol.* **149**, 139-148.
- Rosenbaum, L., Hinselmann, G., Jahn, A. and Zell, A. (2011). Interpreting linear support vector machine models with heat map molecule coloring. *J. Cheminform.* **3**, 11.
- Rothman, J. M., Chapman, C. A. and Van Soest, P. J. (2012). Methods in primate nutritional ecology: a user's guide. *Int. J. Primatol.* **33**, 542-566.
- Sakamoto, K. Q., Sato, K., Ishizuka, M., Watanuki, Y., Takahashi, A., Daunt, F. and Wanless, S. (2009). Can ethograms be automatically generated using body acceleration data from free-ranging birds? *PLoS ONE* **4**, e5379.
- Sato, K., Charrassin, J. B., Bost, C. A. and Naito, Y. (2004). Why do macaroni penguins choose shallow body angles that result in longer descent and ascent durations? *J. Exp. Biol.* **207**, 4057-4065.
- Shawe-Taylor J. and Cristianini, N. (2004). *Kernel Methods for Pattern Analysis*. Cambridge: Cambridge University Press.
- Simeone, A. and Wilson, R. P. (2003). In-depth studies of Magellanic penguin (*Spheniscus magellanicus*) foraging: can we estimate prey consumption by perturbations in the dive profile? *Mar. Biol.* **143**, 825-831.
- Thompson, D. and Fedak, M. A. (2001). How long should a dive last? A simple model of foraging decisions by breath-hold divers in a patchy environment. *Anim. Behav.* **61**, 287-296.
- Viviant, M., Trites, A. W., Rosen, D. A., Monestiez, P. and Guinet, C. (2010). Prey capture attempts can be detected in Steller sea lions and other marine predators using accelerometers. *Polar Biol.* **33**, 713-719.
- Viviant, M., Monestiez, P. and Guinet, C. (2014). Can we predict foraging success in a marine predator from dive patterns only? Validation with prey capture attempt data. *PLoS ONE* **9**, e88503.
- Watanabe, Y. Y. and Takahashi, A. (2013). Linking animal-borne video to accelerometers reveals prey capture variability. *Proc. Natl. Acad. Sci. USA* **110**, 2199-2204.
- Watanabe, S., Izawa, M., Kato, A., Ropert-Coudert, Y. and Naito, Y. (2005). A new technique for monitoring the detailed behaviour of terrestrial animals: a case study with the domestic cat. *Appl. Anim. Behav. Sci.* **94**, 117-131.
- Watanabe, Y. Y., Ito, M. and Takahashi, A. (2014). Testing optimal foraging theory in a penguin-krill system. *Proc. Biol. Sci.* **281**, 20132376.
- Weerheim, M. S., Klomp, N. I., Brunsting, A. M. and Komdeur, J. (2003). Population size, breeding habitat and nest site distribution of little penguins (*Eudyptula minor*) on Montague Island, New South Wales. *Wildl. Res.* **30**, 151-157.
- Wilson, R. P., White, C. R., Quintana, F., Halsey, L. G., Liebsch, N., Martin, G. R. and Butler, P. J. (2006). Moving towards acceleration for estimates of activity-specific metabolic rate in free-living animals: the case of the cormorant. *J. Anim. Ecol.* **75**, 1081-1090.
- Ydesen, K. S., Wisniewska, D. M., Hansen, J. D., Beedholm, K., Johnson, M. and Madsen, P. T. (2014). What a jerk: prey engulfment revealed by high-rate, supercranial accelerometry on a harbour seal (*Phoca vitulina*). *J. Exp. Biol.* **217**, 2239-2243.
- Zimmer, I., Ropert-Coudert, Y., Kato, A., Ancel, A. and Chiaradia, A. (2011a). Does foraging performance change with age in female little penguins (*Eudyptula minor*)? *PLoS ONE* **6**, e16098.
- Zimmer, I., Ropert-Coudert, Y., Poulin, N., Kato, A. and Chiaradia, A. (2011b). Evaluating the relative importance of intrinsic and extrinsic factors on the foraging activity of top predators: a case study on female little penguins. *Mar. Biol.* **158**, 715-722.