Sweeney *et al. Anim Biotelemetry (2019) 7:7*
https://doi.org/10.1186/s40317-019-0169-3

Animal Biotelemetry

RESEARCH

Open Access

# Automated peak detection method for behavioral event identification: detecting *Balaenoptera musculus* and *Grampus griseus* feeding attempts

David A. Sweeney[1*], Stacy L. DeRuiter[2], Ye Joo McNamara-Oh[3], Tiago A. Marques[4,5], Patricia Arranz[6,7] and John Calambokidis[8]

**Abstract**

The desire of animal behaviorists for more flexible methods of conducting inter-study and inter-specific comparisons and meta-analysis of various animal behaviors compelled us to design an automated, animal behavior peak detection method that is potentially generalizable to a wide variety of data types, animals, and behaviors. We detected the times of feeding attempts by 12 Risso's dolphins (*Grampus griseus*) and 36 blue whales (*Balaenoptera musculus*) using the norm-jerk (rate of change of acceleration) time series. The automated peak detection algorithm identified median true-positive rates of 0.881 for blue whale lunges and 0.410 for Risso's dolphin prey capture attempts, with median false-positive rates of 0.096 and 0.007 and median miss rates of 0.113 and 0.314, respectively. Our study demonstrates that our peak detection method is efficient at automatically detecting animal behaviors from multisensor tag data with high accuracy for behaviors that are appropriately characterized by the data time series.

**Keywords:** Blue whale, Detection, Lunge, Norm-jerk, Prey capture, Risso's dolphin

## Background

Studying animal behavior while minimizing levels of invasiveness is a challenge many biologists face [1–3]. Difficulty also arises while attempting to observe animals in environments and during time periods that are relatively inaccessible to humans [4]. The scientific field of biologging arose, in part, to address these two main obstacles [5]. Since the first use of tagging devices on animals in 1963, the field of biologging has evolved into a discipline that allows for the detailed behavioral study of animals ranging from chipmunks to blue whales [5–7].

Over the years, a variety of unique tagging devices (i.e., DTAGs [8], Acousonde tags [9], etc.) have been developed by researchers around the world to try to gain access into the lives of animals. Using data obtained by these tags, scientists can determine the exact time when an animal exhibited a certain behavior [10–17]. For example, a recent study identified potential sleeping behavior in harbor porpoises (*Phocoena phocoena*) by searching for inactive, uniform diving behavior using acceleration, depth, and acoustic data obtained from data-logger tags [14]. Another study used accelerometer tags attached to the heads of two Antarctic penguin species to detect peaks in the acceleration signal and thus study prey encounter rates of these species [15]. Additionally, recent work with Weddell seals (*Leptonychotes weddellii*) and Antarctic fur seals (*Arctocephalus gazella*) used accelerometer tags attached to the mandible to detect signatures in the accelerometer signal indicative of mouth opening associated with feeding events [18, 19]. Due to the often-high sampling rate of multisensor tags and the long duration of data recording, the process of determining the time of every instance of a given behavior can be quite arduous [1].

Scientists have begun to develop automated animal behavior detection algorithms to improve the efficiency

---

*Correspondence: dasweeney4423@gmail.com
[1] Department of Biology, Calvin College, Grand Rapids, MI 49546, USA
Full list of author information is available at the end of the article

of multisensor data analysis. They scan through data searching for signal characteristics that are known to be indicators of unique animal behaviors. Many of these detection algorithms are designed to detect feeding attempts by animals in animal-borne tag data [15, 20–25]. For example, Cox and colleagues developed a method to detect foraging behavior in juvenile southern elephant seals, but their code is highly specific to Argos relay satellite tags and requires triaxial acceleration and pressure data [25]. The detection of feeding attempts is commonly desired, because knowing when a predator hunts (and captures) prey can allow for more in-depth analyses of various ecological parameters, including, for example, studying the overall energy expenditure of foraging species [26].

While many of these event detection methods have very specific data format requirements, more adaptive methods that can easily be applied to data from different species and tag types would facilitate inter-study and inter-specific comparisons and meta-analysis of a variety of different animal behaviors. For a behavior to be detected and analyzed using automated detection methods, a proxy must exist from which the behavior can be identified. Blue whales feed by lunging toward a prey patch while simultaneously opening their mouths, thus generating peaks in the norm-jerk signal due to produced and incurred changes in acceleration [27–29]. Risso's dolphins emit buzz sounds during close pursuit, attempted capture, or capture of prey. Recently published studies have shown that peaks in the norm-jerk signal are often associated with the end of the buzz when the sound is associated with a prey capture attempt [10, 30, 31]. A useful proxy commonly associated with foraging events in cetaceans is therefore the norm-jerk signal.

In this study, we demonstrate the overall performance of a newly designed automated detection method (titled detect_peaks), operational in many versions of three widely used software programs (R [32], MATLAB [33], and Octave [34]), at detecting the times of Risso's dolphin (*Grampus griseus*) and blue whale (*Balaenoptera musculus*) foraging events from the norm-jerk signal. We also compare the accuracy of detections using our method to manual detections and other automated detection methods for two very different species to analyze some of the difficulties involved with this approach.

## Results

The detect_peaks automated detection algorithm detected blue whale and Risso's dolphin feeding attempts with varying accuracies using default and optimized parameters. Default parameters were automatically set by the detect_peaks detection method for each animal, as specified in the methods section. Optimized threshold levels were set as those that returned the best detection results upon receiver operating characteristic (ROC) curve analysis using blanking times of 30 s for the blue whales and 2 s for the Risso's dolphins (i.e., biological blanking times). Only 7 of the 12 Risso's dolphins used in this study had known prey capture events (max = 261, min = 2, median = 2.5, IQR = 51.25). On the other hand, 35 of the 36 blue whales had known times of lunge feeding events (max = 189, min = 1, median = 65.5, IQR = 64). A detection performance summary is presented in Table 1.

For Risso's dolphins, detections using optimized thresholds and biological blanking times returned a median true-positive detection rate (number of true-positive detections divided by the total number of known behavioral events) of 0.410 (IQR = 0.697) and a median false-positive detection rate (number of false-positive detections divided by the total number of possible behavioral events) of 0.007 (IQR = 0.022). The median true-positive rate for optimized blue whale detections (median = 0.881, IQR = 0.136) was better than that of Risso's dolphins. The median false-positive rate for the optimized blue whale detections (median = 0.096, IQR = 0.083) was larger than the Risso's dolphin median false-positive rate (median = 0.007, IQR = 0.022). Optimized blue whale detections produced a median miss rate (number of missed detections divided by the total number of known behavioral events) of 0.113 (IQR = 0.134), which was less than half that of the Risso's dolphin detections (median = 0.314, IQR = 0.154). ROC curves plotting the optimized true-positive rates and false-positive rates for all animals of both species as well as the median rates across all animals are shown in Fig. 1.

Detections using the default parameters for each species seemingly performed better than detections using biological blanking times and optimized thresholds (Table 2). Default parameters for blue whales consisted of a median threshold level of 0.454 (IQR = 0.300) and a median blanking time of 4.700 (IQR = 13.000). The median optimized blue whale threshold was 0.429 (IQR = 0.432). Default parameters for Risso's dolphins consisted of a median threshold level of 3.027 (IQR = 2.724) and a median blanking time of 0.540 (IQR = 0.640). The median optimized Risso's dolphin threshold was 7.774 (IQR = 10.504).

When looking at the side-by-side plots of each dolphin's norm-jerk signal and dive profile (e.g., Fig. 2), we noticed that 95.6% of the known prey capture attempts occurred at depths greater than 10 m. However, there were also strong spikes in the jerk signal, while the dolphin was near the surface, resulting in 1.30% of all false-positive detections occurring while the Risso's dolphins were swimming within 10 m of the water's surface. Upon

**Table 1 Detection performance summary table for both species**

| | | Default detections | | | | Optimal detections | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | **Mean** | **Median** | **SE** | **IQR** | **Mean** | **Median** | **SE** | **IQR** |
| Blue whales | True positives | 58.306 | 55.000 | 6.999 | 54.250 | 56.111 | 49.500 | 6.693 | 54.750 |
| | False positives | 222.528 | 127.000 | 40.895 | 102.250 | 57.361 | 45.000 | 7.453 | 59.250 |
| | Misses | 6.886 | 4.000 | 1.148 | 11.000 | 9.143 | 7.000 | 1.326 | 10.500 |
| | True-positive rate | 0.865 | 0.918 | 0.031 | 0.132 | 0.835 | 0.881 | 0.028 | 0.136 |
| | False-positive rate | 0.067 | 0.034 | 0.010 | 0.081 | 0.104 | 0.096 | 0.010 | 0.083 |
| | Misses rate | 0.110 | 0.079 | 0.019 | 0.122 | 0.141 | 0.113 | 0.015 | 0.134 |
| | Distance from [0,1] | 0.162 | 0.128 | 0.031 | 0.161 | 0.205 | 0.172 | 0.028 | 0.150 |
| Risso's dolphins | True positives | 27.833 | 2.500 | 15.345 | 37.500 | 26.000 | 2.000 | 14.624 | 30.500 |
| | False positives | 606.000 | 218.500 | 312.347 | 243.250 | 85.167 | 31.500 | 33.049 | 145.500 |
| | Misses | 23.714 | 14.000 | 11.004 | 33.000 | 26.857 | 13.000 | 12.160 | 36.500 |
| | True-positive rate | 0.441 | 0.492 | 0.122 | 0.760 | 0.385 | 0.410 | 0.108 | 0.697 |
| | False-positive rate | 0.053 | 0.015 | 0.027 | 0.022 | 0.017 | 0.007 | 0.007 | 0.022 |
| | Misses rate | 0.243 | 0.279 | 0.084 | 0.286 | 0.339 | 0.314 | 0.080 | 0.154 |
| | Distance from [0,1] | 0.568 | 0.508 | 0.122 | 0.759 | 0.617 | 0.593 | 0.107 | 0.695 |

The table shows the mean and median values representing the overall performance of the detection method across all individuals within each species as well as the standard error (SE) and interquartile range (IQR) for each element of the table. "Distance from [0,1]" refers to ROC curve analysis, with smaller values (closer to 0) representing better performance. Note that the performance statistics for each individual animal in this study are shown in Additional file 1

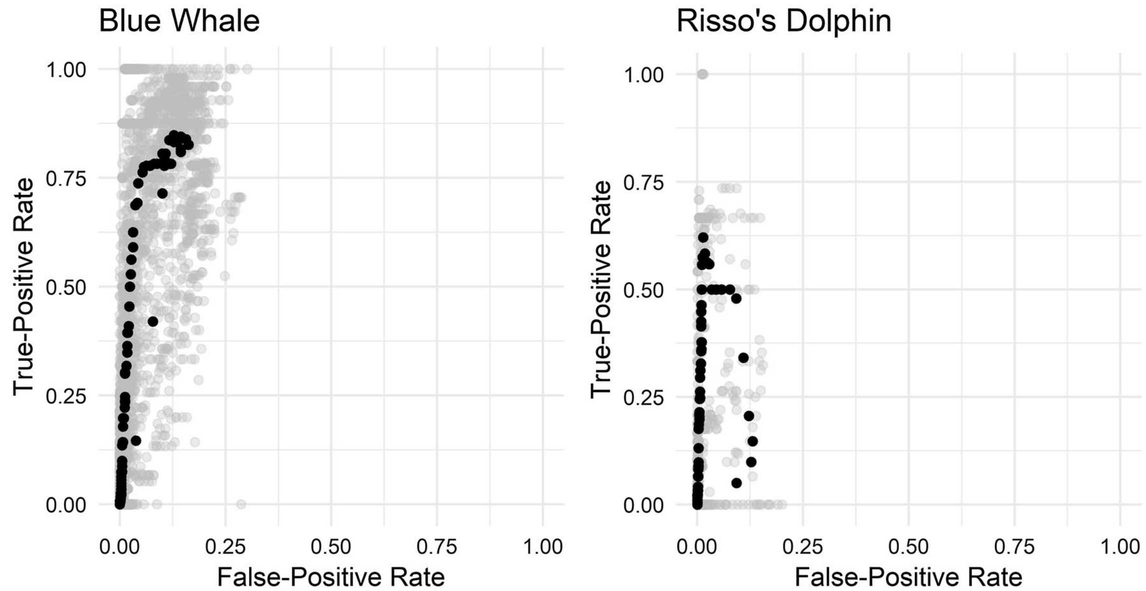


**Fig. 1** ROC curves showing true-positive versus false-positive rates and median rates across all animals. The median detection rates for each species are shown in black, while the gray points display the overall spread of the detections for each animal at every threshold level. The gray points are semitransparent, so areas of darker shading represent areas with more points

conducting detections after removing all data when the dolphins were shallower than 10 m, we observed that the results from these detections did not produce drastically better false-positive detection rates (mean false-positive rate improved by 0.001 and median false-positive rate improved by 0.004) and our true-positive detection rates

**Table 2 Table of detection parameters used in this study**

| | Blue whales | | | | Risso's dolphins | | | |
|---|---|---|---|---|---|---|---|---|
| | Mean | Median | SE | IQR | Mean | Median | SE | IQR |
| Default threshold level | 0.583 | 0.454 | 0.064 | 0.300 | 18.933 | 3.027 | 15.864 | 2.724 |
| Default blanking time | 8.584 | 4.700 | 1.341 | 13.000 | 2.368 | 0.540 | 1.340 | 0.640 |
| Optimal threshold level | 0.726 | 0.429 | 0.149 | 0.432 | 19.276 | 7.774 | 8.486 | 10.504 |
| Biological blanking time | 30.000 | 30.000 | 0.000 | 0.000 | 2.000 | 2.000 | 0.000 | 0.000 |

Threshold levels and blanking times used while performing the default and optimized detections are also shown. Note that the threshold levels and blanking times used for each specific individual are shown in Additional file 1

decreased for some individuals. We, therefore, included data from all depths in our analysis. Risso's dolphin prey capture attempts were best-detected when the associated peak was at least above the 0.9 quantile of the norm-jerk signal. Peaks below this level were still frequently detected, but the probability of these detections resulting in a missed detection gradually increased as the threshold level decreased.

Blue whale side-by-side plots (e.g., Fig. 2) showed that many false-positive detections occurred at times while the whale was at or near the surface of the water. 90.6% of lunges occurred at depths greater than 10 m. The roughly 9.4% of lunges that occurred near the water's surface were detected 52.0% of the time. In contrast to the often-sporadic peaks (large peaks associated with non-foraging, unknown behaviors) in the Risso's dolphin norm-jerk signals, blue whales' norm-jerk signals appear to have much more uniformity. For many of the blue whales, the strongest jerk signal is during a foraging event, with fewer occasions of abnormally strong peaks representing behaviors other than feeding attempts compared to Risso's dolphins (see Additional files 2 and 3 for norm-jerk signals with marked prey captures and lunges for all animals).

## Discussion

We have developed an automated behavioral event detection method, which is successful at identifying the times of blue whale and Risso's dolphin feeding attempts using the norm-jerk time series. The accuracy of the detections does vary, however, between species and across individuals. We observed that the norm-jerk signal is not as good of a proxy for detecting feeding attempts for Risso's dolphins as it is for blue whales.

The success of the blue whale detections seems to be due to the tendency for the largest peaks in the norm-jerk signals to be representative of lunges. This allowed for more accurate detections with fewer false positives and misses. These large peaks are caused by dramatic deceleration of blue whales during feeding lunges where opening of the mouth and filling of the buccal pouch create a sharp increase in drag [28]. This large ratio of prey capture jerk peaks to overall signal noise is likely due to the large overall body mass of blue whales. Cetaceans with greater body mass have been shown to exhibit lower overall stroke frequencies, consequentially minimizing the norm-jerk signal noise at times when the whale is traveling at relatively constant rates [35].

For both species, default detections generally performed better than the optimized parameter detections according to ROC curve analyses. Although the default thresholds were relatively similar to the optimized thresholds, the default blanking times were generally far lower than the biologically predetermined blanking times. Blanking times are used to reduce the number of false-positive detections by allowing for multiple signal values to be considered one animal behavior. Therefore, the biological blanking times were set based on previous research that discusses the durations of the desired behaviors. A lower blanking time allows for more detection to be made, thereby commonly increasing the total number of true-positive detections. The number of false-positive detections also increases with lower blanking times, but due to the extremely large total number of possible behavior events, the false-positive rate increases at a drastically lower rate than does the true-positive rate per detection made. From this observation, we recommend that future users of this behavioral event detection method should not fret too much over setting the "perfect/optimal" threshold level and blanking time. We highly encourage all future users of this automated detection method to perform post hoc analyses of the events detected, given that no matter what parameters are used to perform the detections, it is highly unlikely that a true-positive rate of 1.0 and false-positive rate of 0.0 will be obtained.

Risso's dolphin detections contained many false-positive detections and missed detections across individuals. Risso's dolphin detections also have very low true-positive detection rates. The large number of false-positive detections and somewhat low number of true-positive detections seem to beget the conclusion that the norm-jerk is not an effective proxy for detecting prey capture attempts for this species. The number of desired peaks

(See figure on next page.)

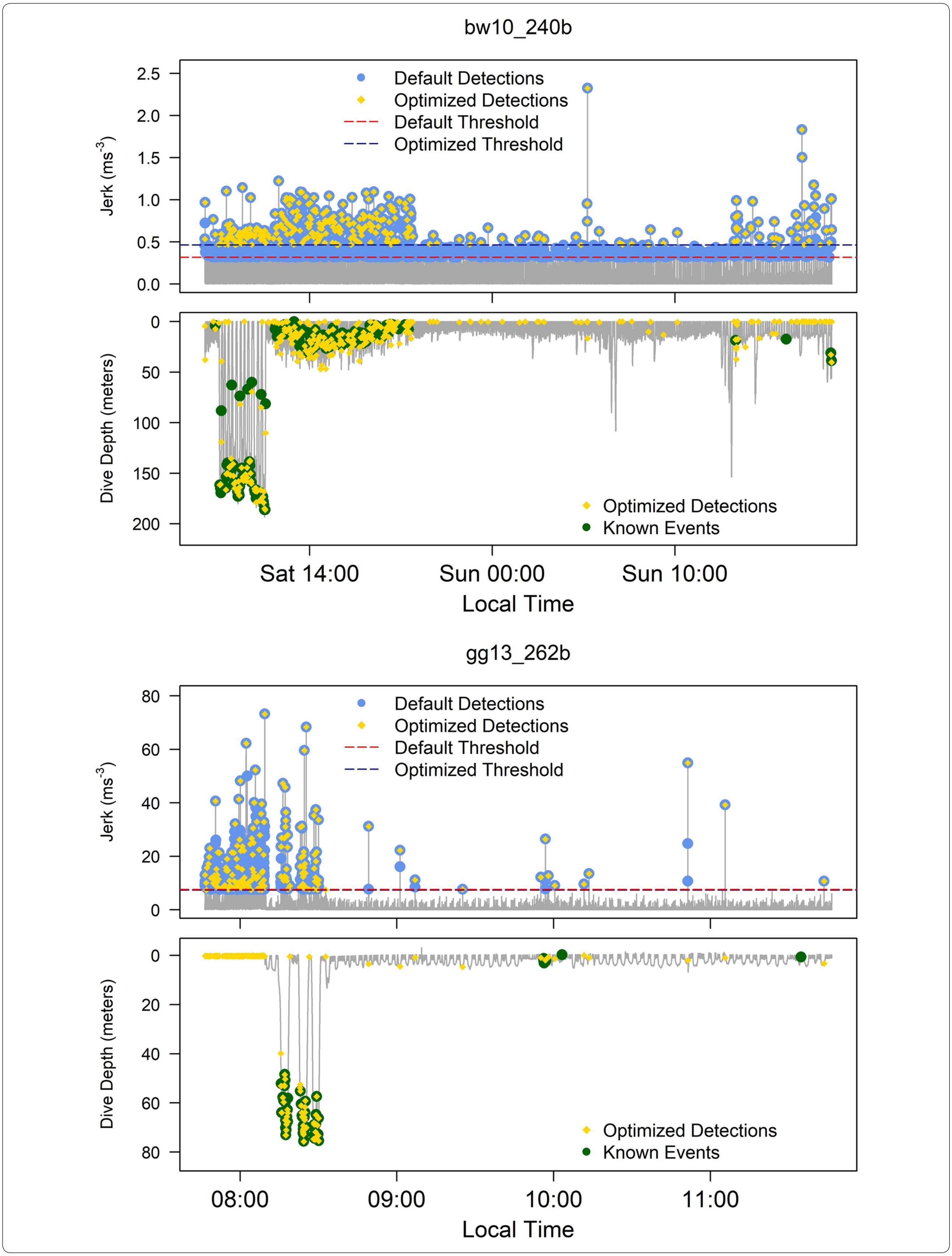
**Fig. 2** Plots of norm-jerk signals and dive depth for one blue whale (bw10_240b) and one Risso's dolphin (gg13_262b). Plots on the top for each individual show the norm-jerk signals that were passed through detect_peaks. Default and optimized detections are labeled with their corresponding threshold levels. The bottom plots for each individual represent the dive depth of each animal with known feeding attempts and the optimized detections marked on the plot. Note that the optimized and default thresholds for the Risso's dolphin are almost identical, thus seemingly overlapping in the figure. The sampling rates for this blue whale and Risso's dolphin were 5 Hz and 25 Hz, respectively

representing prey capture events is too similar to the number of undesired peaks representing other common Risso's dolphin behaviors (e.g., playful socializing and energetic traveling [36]), thereby bringing about a large number of false-positive detections for this species. Some of the missed detections we presume may be due to prey capture attempts when the dolphin did not have to maneuver rapidly to catch a potentially stationary prey. Similarly, the magnitude of peaks during prey capture attempts could differ depending on the DTAG's (suction-cup-attached digital tag) location on the dolphin. Tag placement can vary between animals due to the difficulties of attaching suction-cup tags or the possibility that a suction-cup tag could slide while recording [37]. These changes in tag placement can affect accelerometer signals, thus altering the norm-jerk and potentially leading to differences in overall detection rates [37]. There is also the possibility that missed detections were due to buzzes in which the animal aborted the prey capture attempt, buzzes made in a social context, or buzzes that were produced by a nearby conspecific [30].

When comparing the overall performance of our automated detection method against those previously developed, we observe that our detection method performed similarly despite the intentional simplicity of our detection algorithm. Not every paper mentioned earlier that describes an automated method to detect prey captures has listed accuracy statistics. Some, however, do have detection accuracy statistics. Owen et al. obtained a true-positive detection rate of approximately 0.700 for the known lunges of humpback whale (*Megaptera novaeangliae*) surface, lunge feeding events using a combination of acceleration and pitch data from DTAGs. They obtained a false-positive detection rate of roughly 0.200 [24]. Allen et al. obtained a true-positive detection rate of approximately 0.920 for fin whale lunge feeding events (a species with similar lunge feeding to blue whales) using a decision-tree method that incorporated a combination of jerk, depth, roll, and flow noise data from DTAGs. They obtained a false-positive detection rate of roughly 0.310 [23]. In comparison, our detection statistics show that we obtained a median true-positive detection of 0.881 for the known lunges of blue whales, with a median false-positive rate of 0.096.

Among previously developed automated approaches to detect feeding attempts by species other than rorqual whales, Viviant et al.'s optimal method obtained a true-positive detection rate of about 0.90 for the known feeding attempts of Steller sea lions (*Eumetopias jubatus),* with a false-positive rate of about 0.25 using accelerometer data from Little Leonardo acceleration data loggers [15]. Cox et al. had true-positive detection rates of about 0.59 for juvenile southern elephant seals (*Mirounga leonina*), with false-positive detection rates of about 0.02 using a combination of depth, satellite, acceleration, and pitch data from custom-designed Argos relay satellite tags [25]. Although in no way identical, these seal feeding attempt detection methods are perhaps more closely related to our Risso's dolphins than our blue whales due to the enhanced maneuverability of seals and dolphins compared to rorquals. One caveat worth mentioning is that the tags in the Cox and Viviant studies were attached to the heads of the seals, whereas the tags on the Risso's dolphins were initially attached near the dorsal fins. A tag attached to the head of a seal would record changes in acceleration due to both total body acceleration and potential head maneuvering while foraging. Conversely, tags attached near a dorsal fin would predominantly record changes in total body acceleration. Also, many cetaceans (including Risso's dolphins) have fused cervical vertebrae, thus severely minimizing head maneuvering. Another caveat to be considered is prey-type preferences for these species. Risso's dolphins often have different prey preferences compared to southern elephant seals and Steller sea lions, likely resulting in different accelerometer signatures during prey catches. These resulting differences in accelerometer signatures could influence the accuracy of different detection algorithms. That being said, our Risso's dolphin optimized detection algorithm returned a median true-positive rate of 0.410, with a median false-positive rate of 0.007.

In making detect_peaks, we created a peak detection method that allows for the generalized automated detection of any behavioral event, given that the signal input to the detection algorithm is a good proxy for predicting the specific behavior. Ideas for expanding upon our detection method's current design have been proposed: allowing for bivariate detections, incorporating an additional parameter to adjust a maximum behavioral event duration, and integrating time-varying parameters. However, these ideas were not implemented in the current algorithm because the goal of designing detect_peaks was to



Content courtesy of Springer Nature, terms of use apply. Rights reserved.

create an easy-to-use, efficient, and flexible behavioral event detection method. We currently feel that expanding on the current design of the detection method would infringe upon this goal.

Based on the results of our foraging event detections, it appears that the norm-jerk signal is a good proxy for detecting blue whale lunges and a good, although somewhat less effective, proxy for detecting Risso's dolphin prey capture attempts. More research may help identify a better input signal for detect_peaks (one that has strong peaks in the signal only during feeding attempts) to allow for the enhanced detection of Risso's dolphin prey captures. Future work could also shed light on how to best utilize blanking times in the detect_peaks algorithm and improve the precision of biologically predetermined blanking times for animal behaviors.

## Conclusions

The performance characteristics of detect_peaks alone show evidence for the usefulness of this automated behavioral event detection algorithm, given that they perform at similarly high levels compared to previously developed methods. However, unlike the other previously mentioned detection methods [20, 23–25], detect_peaks was intentionally designed to be capable of detecting a potentially endless list of behaviors from many different species. The simple algorithm used by detect_peaks has potential for use in real-time or on-board processing in telemetry tags, if validated for a particular species and tag type. In addition, given that many scientists have limited time or software development capabilities, we believe that making this detection method freely available as part of open-source software for high-resolution movement-sensing tags has the potential to make event detection in biologging data easier and more reproducible.

## Methods

### Data collection and preparation

This project utilized data from suction-cup digital acoustic recording tags (DTAGs) attached to 36 blue whales and 12 Risso's dolphins. Each tag recorded acoustic data using hydrophones and recorded animal movement data using pressure sensors, triaxial accelerometers, and magnetometers [8]. The 36 blue whales were tagged between 2010 and 2013 in and around the Southern California Bight by members of the Southern California Behavioral Response Study (SOCAL BRS), and movement sensors were sampled at 5 to 25 Hz [38]. The 12 Risso's dolphins were tagged in 2011, 2013, and 2014 mostly around Catalina Island off the coast of California, USA, and movement sensors were sampled at 10 to 200 Hz [30]. All data were obtained in accordance with the US National Marine Fisheries Service permits #14534 and #19116.

Data obtained from each blue whale's tag were cropped to remove samples of times when the tag was not attached to the whale. The Risso's dolphin data were further cropped for consistency with previous studies, removing the first fifteen minutes of tag recording to exclude data potentially influenced by the tagging procedure, and also removing data recorded after the beginning of controlled acoustic exposure experiments or data recorded after the tag had already fallen off the animal [30]. All analyses were performed in R [32] and MATLAB [33] using functions from the tagtools package (https://github.com/stacyderuiter/TagTools).

### Feeding attempt detections

The times of cetacean foraging events have been previously determined using kinematic data obtained from animal–borne tags [22–24]. A time series commonly used in the identification of foraging events is the norm-jerk signal, which at time $t$ is represented by:

$$j_t = \|A_t - A_{t+1}\| * S$$

$A$ is the triaxial acceleration matrix at time $t$, and $S$ is the sampling rate. Rorqual lunge feeding events exhibit large peaks in the norm-jerk signal due to the sudden changes in acceleration related to the increased speed upon approach of a prey patch and the drastic decrease in acceleration caused by induced drag upon opening of the mouth [27, 29, 39]. Similarly, an association has been observed between strong jerk signals and buzzes, which are known to commonly represent prey capture attempts, in several odontocete species due to the rapid physical maneuvering required to catch prey items [10, 30, 40].

We have hence developed an automated behavioral event detection algorithm that operates as a threshold detection method where peaks that surpass a specified threshold level are labeled as the behavioral event of interest. The detection method, titled detect_peaks, is generalizable to a wide variety of potential data types, animals, and behaviors. Detect_peaks allows for the input of any type of time series or a matrix accompanied by a separate function that converts the matrix into a time series. The time series that is used by detect_peaks may contain positive and/or negative values. For the best detection results, the time series should have spikes (larger values) coinciding in time with the behavioral event and small values otherwise.

Upon running the detect_peaks algorithm, we computed the norm-jerk from the animal's triaxial acceleration. Then, we marked all samples in the norm-jerk signal that surpassed a user-adjustable threshold level as candidate behavioral events. All peaks that surpassed the threshold level were then broken up into individual behavioral events using the blanking time, which is also user-adjustable. The blanking time is a specified length of time between signal peaks detected above the threshold level (from the moment the first peak recedes below the threshold level to the moment the second peak surpasses the threshold level again). If the time between peaks is greater than the specified blanking time, each peak is labeled as a unique behavioral event. If the time between peaks is less than the specified blanking time, the two peaks are grouped into one larger behavioral event. Blanking times are used to account for physical and physiological restrictions upon the minimum possible time between feeding attempts (or other behavioral events). The time at which the maximum norm-jerk level was reached for each behavioral event and the start and end times of the behavioral event were obtained upon completion of the detections.

Known lunge times for the blue whales were determined by expert human analysts who looked for characteristic patterns associated with lunge feeding in a combination of plots consisting of the animal's acceleration, dive depth, body orientation, and swim speed [11, 12, 27, 28, 41]. For the blue whales, a detection was considered a true positive if it was found to exist within a 10 s window (5 s before and 5 s after) of the known lunge time. If the detection was outside of this time window, it was counted as a false-positive detection. This time window was used to account for the possible differences between the times at which the maximum norm-jerk level was reached and the SOCAL BRS members' manual detection times. Given that lunge behaviors for large rorquals are known to last approximately 15 s and are often separated by 30 s of time to allow for proper water filtration and for the whale to travel to a new prey patch, it is highly unlikely that this size time window has caused biased detection results [11]. For the Risso's dolphins, we used the times of buzzes as the known times of prey capture attempts. Buzzes are rapid echolocation click series that are commonly interpreted as attempts to capture prey [10, 30, 42, 43]. Buzz times were first determined by aural and visual inspection of spectrograms by Arranz et al. [30]. Arranz et al. then generated a multivariate Gaussian mixture model that distinguished buzzes from other communication-related pulsed sounds on the basis of duration, temporal proximity to regular echolocation clicks, and jerk ratios [30]. The Risso's dolphin's detections were true positives if they occurred within a 4-s window (2 s before and 2 s after) of the time of the end of known buzz times. If the detection was outside of this time window, it was counted as a false-positive detection. A 4 s window was used to account for the occasions when the maximum peak of the norm-jerk signal did not line up precisely with the end of the buzz sequence. For both species, if an instance of a known behavioral event was not detected, it was counted as a missed detection.

Detections were performed on each animal twice: once using detect_peaks's default threshold and blanking time parameters and once using biologically predetermined blanking times and threshold levels. The default threshold is set as the 0.99 quantile of the norm-jerk signal, and the default blanking time is set as the 0.80 quantile of the time differences between consecutive signal values that surpass the threshold level. The biological blanking times for each species were determined based on previously published observations on feeding behaviors. The biological blanking time for the Risso's dolphins was set at 2 s (a conservative estimate based on observed buzz durations of about 1 s in related species: false killer whales (*Pseudorca crassidens*) and bottlenose dolphins (*Tursiops truncatus*) [31]). The biological blanking time for the blue whale detections was set to 30 s, because a previous study on a group of fin whales (*Balaenoptera physalus*) observed the minimum time between consecutive lunge feeding events to be around 30 s [11].

Optimal thresholds were determined for each individual using receiver operating characteristic (ROC) curves (all ROC curves are available in Additional files 2 and 3) ROC curves were constructed for each individual by running detect_peaks one hundred times using the biological blanking time for that species and one hundred different threshold levels. These different thresholds were equally spaced starting at one hundredth of the maximum norm-jerk signal value for the individual and going to the maximum norm-jerk signal value. True-positive and false-positive rates were calculated for each threshold, and they were all plotted to form the ROC curve. True-positive rates were calculated as the number of true-positive detections divided by the total number of known behavioral events. False-positive rates were calculated as the number of false-positive detections divided by the total number of possible behavioral events (set as the duration of the tag recording (in seconds) divided by the blanking time with the total number of known behavioral events subtracted).

After the ROC curve was completed for each individual, the optimal threshold was set as that which produced true-positive and false-positive rates closest to the upper-left corner of the plot (corresponding to a true-positive rate of one and a false-positive rate of zero). Although we defined this as the "optimal" threshold level, different instances in different studies may prefer the "optimal" threshold level to be determined based on a different set of criteria. However, for the sake of maintaining consistency, we will refer to all detection results using the threshold level as determined by our ROC curve criteria mentioned previously as the optimal detection results. Threshold optimizations were performed for the purpose of testing the effectiveness of our biological blanking times with threshold levels that were determined to return accurate detections. Automated detections using the detect_peaks's default settings were performed for all animals, including those without known prey capture attempts. For the animals that did not have any known feeding attempts, their default and optimized true-positive rates were always zero. The optimal threshold for these animals was set as the highest value of the norm-jerk signal, because this threshold always produced the absolute minimum false-positive rate; the threshold level allowed for only one false-positive detection.

Analysis on the overall performance of the detection method was done by comparing the performance statistics as listed in Table 1 with those of other, previously published, automated behavioral event detection algorithms. Additional analyses were done by creating side-by-side plots (e.g., Fig. 2) of the norm-jerk signal and the dive profile of each individual and then observing trends in predation behaviors and trends in the detections made by the detect_peaks algorithm. For the Risso's dolphins, an additional set of detections was performed with all jerk peaks, while the animal was within 10 m of the water's surface removed. This was done in an effort to decrease the false-positive rate of the Risso's dolphin detections. However, because these detection results were not drastically better, we included data from all depths in our final analyses.

## Additional files

**Additional file 1.** Data spreadsheets displaying the detection statistics and tag deployment information for all animal individuals used in this study.

**Additional file 2.** File containing ROC curve and side-by-side plot figures similar to those displayed above in the manuscript for all Risso's dolphin individuals used in this study.

**Additional file 3.** File containing ROC curve figures and side-by-side plot figures similar to those displayed in the manuscript for all blue whale individuals used in this study.

### Authors' contributions

DAS wrote the manuscript. SLD, YJM, TAM, PA, and JC assisted with manuscript preparation. DAS, SLD, and YJM developed detect_peaks. DAS performed all analyses. DAS, SLD, TAM, and YJM assisted in creating the tagtools kit, within which detect_peaks is included. PA and JC provided data for the project. SLD obtained funding for the project. All authors read and approved the final manuscript.

### Author details

[1] Department of Biology, Calvin College, Grand Rapids, MI 49546, USA. [2] Department of Mathematics and Statistics, Calvin College, Grand Rapids, MI 49546, USA. [3] Department of Computer Science, Calvin College, Grand Rapids, MI 49546, USA. [4] Centre for Research into Ecological and Environmental Modelling, The Observatory, University of St Andrews, St Andrews KY16 9LZ, UK. [5] Departmamento de Biologia Animal, Centro de Estatística e Aplicações, Faculdade de Ciências, Universidade de Lisboa, 1749-016 Lisbon, Portugal. [6] Department of Animal Biology, Universidad de La Laguna, Avd. Astrofisico Fco Sanchez s/n, 36200 La Laguna, Tenerife, Spain. [7] Sea Mammal Research Unit, School of Biology, University of St Andrews, East Sands, St Andrews KY16 8LB, UK. [8] Cascadia Research Collective, Olympia, WA 98501, USA.

### Acknowledgements

We acknowledge the leaders and participants of the August 2017 Animal Tag Tools Workshop for their helpful insights into improving this peak detection method in its early stages of development. We also wish to acknowledge all members of the SOCAL-Behavioral Response Study project for their work obtaining and sharing data for this study. Detect_peaks was developed and distributed as part of a set of software tools for high-resolution movement-sensing tags (https://github.com/stacyderuiter/TagTools). Help documentation and further information about the tagtools project can be found at our website (http://www.animaltags.org/doku.php).

### Competing interests

The authors declare that they have no competing interests.

### Availability of data and materials

The datasets used during the current study and supporting information are available from the Dryad Digital Repository: https://datadryad.org/review?doi=doi:10.5061/dryad.bd8j403.

### Consent for publication

Not applicable.

### Ethics approval and consent to participate

Experiments with marine mammals were carried out under U.S. National Marine Fisheries Service permits #14534 and #19116. Experimental protocols were approved by the IACUC at Cascadia Research Collective.

### Funding

This project was supported by the U.S. Office of Naval Research (Grant No. N00014-16-1-3089). Tag deployments used here from the SOCAL-Behavioral Response Study were funded by the U.S. Office of Naval Research and Navy Living Marine Resources under multiple grants and contracts. Further funding was provided by the Kuipers Research Fellowship and the Calvin Alumni Association. TAM thanks partial support by CEAUL (funded by FCT—Fundação para a Ciência e a Tecnologia, Portugal, through the project UID/MAT/00006/2013).

## Publisher's Note

Received: 8 November 2018 Accepted: 19 March 2019

### References

1. Brown DD, Kays R, Wikelski M, Wilson R, Klimley AP. Observing the unwatchable through acceleration logging of animal behavior. Anim Biotelemetry. 2013;1:1–16.
2. Schneirla TC. The relationship between observation and experimentation in the field study of behavior. Ann N Y Acad Sci. 1950;51:1022–44.
3. van der Hoop JM, Fahlman A, Shorter KA, Gabaldon J, Rocho-Levine J, Petrov V, et al. Swimming energy economy in bottlenose dolphins under variable drag loading. Front Mar Sci. 2018;5:465. https://doi.org/10.3389/fmars.2018.00465.
4. Joyce TW, Durban JW, Claridge DE, Dunn CA, Fearnbach H, Parsons KM, et al. Physiological, morphological, and ecological tradeoffs influence

vertical habitat use of deep-diving toothed-whales in the Bahamas. PLoS ONE. 2017;12:1–27.

5. Kooyman GL. Genesis and evolution of bio-logging devices: 1963–2002. Mem Natl Polar Res Inst. 2004;58:15–22.
6. Hammond TT, Springthorpe D, Walsh RE, Berg-Kirkpatrick T. Using accelerometers to remotely and automatically characterize behavior in small animals. J Exp Biol. 2016;219:1618–24. https://doi.org/10.1242/jeb.136135.
7. Acevedo-Gutierrez A, Croll DA, Tershy BR. High feeding costs limit dive time in the largest whales. J Exp Biol. 2002;205:1747–53.
8. Johnson MP, Tyack PL. A digital acoustic recording tag for measuring the response of wild marine mammals to sound. IEEE J Ocean Eng. 2003;28:3–12.
9. Burgess WC, Tyack PL, Le Boeuf BJ, Costa DP. A programmable acoustic tag and first results from free ranging northern elephant seals. Deep Res Part II Top Stud Oceanogr. 1998;45:1327–51.
10. Miller PJO, Johnson MP, Tyack PL. Sperm whale behaviour indicates the use of echolocation click buzzes "creaks" in prey capture. Proc R Soc B Biol Sci. 2004;271:2239–47. https://doi.org/10.1098/rspb.2004.2863.
11. Goldbogen JA, Calambokidis J, Shadwick RE, Oleson EM, McDonald MA, Hildebrand JA. Kinematics of foraging dives and lunge-feeding in fin whales. J Exp Biol. 2006;209:1231–44. https://doi.org/10.1242/jeb.02135.
12. Goldbogen JA, Calambokidis J, Croll DA, Harvey JT, Newton KM, Oleson EM, et al. Foraging behavior of humpback whales: kinematic and respiratory patterns suggest a high cost for a lunge. J Exp Biol. 2008;211:3712–9. https://doi.org/10.1242/jeb.023366.
13. Johnson M, De Soto NA, Madsen PT. Studying the behaviour and sensory ecology of marine mammals using acoustic recording tags: a review. Mar Ecol Prog Ser. 2009;395:55–73.
14. Wright AJ, Akamatsu T, Mouritsen KN, Sveegaard S, Dietz R, Teilmann J. Silent porpoise: potential sleeping behaviour identified in wild harbour porpoises. Anim Behav. 2017;133:211–22.
15. Kokubun N, Kim J-H, Shin H-C, Naito Y, Takahashi A. Penguin head movement detected using small accelerometers: a proxy of prey encounter rate. J Exp Biol. 2011;214:3760–7. https://doi.org/10.1242/jeb.058263.
16. Hernández-Pliego J, Rodríguez C, Dell'Omo G, Bustamante J. Combined use of tri-axial accelerometers and GPS reveals the flexible foraging strategy of a bird in relation to weather conditions. PLoS ONE. 2017;12:e0177892.
17. Wang Y, Nickel B, Rutishauser M, Bryce CM, Williams TM, Elkaim G, et al. Movement, resting, and attack behaviors of wild pumas are revealed by tri-axial accelerometer measurements. Mov Ecol. 2015;3:2.
18. Iwata T, Sakamoto KQ, Takahashi A, Edwards EWJ, Staniland IJ, Trathan PN, et al. Using a mandible accelerometer to study fine-scale foraging behavior of free-ranging Antarctic fur seals. Mar Mammal Sci. 2012;28:345–57.
19. Naito Y, Bornemann H, Takahashi A, McIntyre T, Plötz J. Fine-scale feeding behavior of Weddell seals revealed by a mandible accelerometer. Polar Sci. 2010;4:309–16.
20. Viviant M, Trites AW, Rosen DAS, Monestiez P, Guinet C. Prey capture attempts can be detected in Steller sea lions and other marine predators using accelerometers. Polar Biol. 2010;33:713–9.
21. Doniol-Valcroze T, Lesage V, Giard J, Michaud R. Optimal foraging theory predicts diving and feeding strategies of the largest marine predator. Behav Ecol. 2011;22:880–8.
22. Ware C, Friedlaender AS, Nowacek DP. Shallow and deep lunge feeding of humpback whales in fjords of the West Antarctic Peninsula. Mar Mammal Sci. 2011;27:587–605.
23. Allen AN, Goldbogen JA, Friedlaender AS, Calambokidis J. Development of an automated method of detecting stereotyped feeding events in multisensor data from tagged rorqual whales. Ecol Evol. 2016;6:7522–35.
24. Owen K, Dunlop RA, Monty JP, Chung D, Noad MJ, Donnelly D, et al. Detecting surface-feeding behavior by rorqual whales in accelerometer data. Mar Mammal Sci. 2016;32:327–48.
25. Cox SL, Orgeret F, Gesta M, Rodde C, Heizer I, Weimerskirch H, et al. Processing of acceleration and dive data on-board satellite relay tags to investigate diving and foraging behaviour in free-ranging marine predators. Methods Ecol Evol. 2018;9:64–77.
26. Stephens DW, Krebs JR. Foraging theory. Princeton: Princeton University Press; 1986.
27. Goldbogen JA, Pyenson ND, Shadwick RE. Big gulps require high drag for fin whale lunge feeding. Mar Ecol Prog Ser. 2007;349:289–301.
28. Goldbogen JA, Calambokidis J, Oleson E, Potvin J, Pyenson ND, Schorr G, et al. Mechanics, hydrodynamics and energetics of blue whale lunge feeding: efficiency dependence on krill density. J Exp Biol. 2011;214:131–46. https://doi.org/10.1242/jeb.048157.
29. Simon M, Johnson M, Madsen PT. Keeping momentum with a mouthful of water: behavior and kinematics of humpback whale lunge feeding. J Exp Biol. 2012;215:3786–98. https://doi.org/10.1242/jeb.071092.
30. Arranz P, DeRuiter SL, Stimpert AK, Neves S, Friedlaender AS, Goldbogen JA, et al. Discrimination of fast click-series produced by tagged Risso's dolphins (*Grampus griseus*) for echolocation or communication. J Exp Biol. 2016;219:2898–907. https://doi.org/10.1242/jeb.144295.
31. Wisniewska DM, Johnson M, Nachtigall PE, Madsen PT. Buzzing during biosonar-based interception of prey in the delphinids *Tursiops truncatus* and *Pseudorca crassidens*. J Exp Biol. 2014;217:4279–82.
32. R Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. 2017. https://www.R-project.org/. Accessed 27 Mar 2019.
33. The Mathworks Inc. Matlab Release 2017a. Natick, Massachusetts, United States. 2017.
34. Eaton JW, Bateman D, Hauberg S, Wehbring R. GNU Octave version 4.2.1 manual: A high-level interactive language for numerical computations. 2017. https://www.gnu.org/software/octave/doc/v4.2.1/. Accessed 27 Mar 2019.
35. Sato K, Watanuki Y, Takahashi A, Miller PJ, Tanaka H, Kawabe R, et al. Stroke frequency, but not swimming speed, is related to body size in free-ranging seabirds, pinnipeds and cetaceans. Proc R Soc B Biol Sci. 2007;274:471–7. https://doi.org/10.1098/rspb.2006.0005.
36. Pereira JNDSG. Field notes on Risso's dolphin (*Grampus griseus*) distribution, social ecology, behaviour, and occurrence in the Azores. Aquat Mamm. 2008;34:426–35.
37. Martín López LM, Miller PJO, Aguilar de Soto N, Johnson M. Gait switches in deep-diving beaked whales: biomechanical strategies for long-duration dives. J Exp Biol. 2015;218:1325–38.
38. Southall BL, Moretti D, Abraham B, Calambokidis J, DeRuiter SL, Tyack PL. Marine mammal behavioral response studies in Southern California: advances in technology and experimental methods. Mar Technol Soc J. 2012;46:48–59. https://doi.org/10.4031/MTSJ.46.4.1.
39. Owen K, Jenner CS, Jenner MNM, Andrews RD. A week in the life of a pygmy blue whale: migratory dive depth overlaps with large vessel drafts. Anim Biotelemetry. 2016;4:1–11.
40. Aguilar de Soto N, Madsen PT, Tyack P, Arranz P, Marrero J, Fais A, et al. No shallow talk: Cryptic strategy in the vocal communication of Blainville's beaked whales. Mar Mammal Sci. 2012;28:75–92.
41. Calambokidis J, Schorr GS, Steiger GH, Francis J, Bakhtiari M, Marshall G, et al. Insights into the underwater diving, feeding, and calling behavior of blue whales from a suction-cup-attached video-imaging tag (Crittercam). Mar Technol Soc J. 2007;41:19–29.
42. Johnson M, Madsen PT, Zimmer WMX, de Soto NA, Tyack PL. Foraging Blainville's beaked whales (*Mesoplodon densirostris*) produce distinct click types matched to different phases of echolocation. J Exp Biol. 2006;209:5038–50. https://doi.org/10.1242/jeb.02596.
43. Miller LA, Pristed J, Møhl B, Surlykke A. The click-sounds of narwhals (*Monodon monoceros*) in Inglefield Bay, Northwest Greenland. Mar Mammal Sci. 1995;11:491–502.

# Terms and Conditions

Springer Nature journal content, brought to you courtesy of Springer Nature Customer Service Center GmbH (“Springer Nature”).

Springer Nature supports a reasonable amount of sharing of research papers by authors, subscribers and authorised users (“Users”), for small-scale personal, non-commercial use provided that all copyright, trade and service marks and other proprietary notices are maintained. By accessing, sharing, receiving or otherwise using the Springer Nature journal content you agree to these terms of use (“Terms”). For these purposes, Springer Nature considers academic use (by researchers and students) to be non-commercial.

These Terms are supplementary and will apply in addition to any applicable website terms and conditions, a relevant site licence or a personal subscription. These Terms will prevail over any conflict or ambiguity with regards to the relevant terms, a site licence or a personal subscription (to the extent of the conflict or ambiguity only). For Creative Commons-licensed articles, the terms of the Creative Commons license used will apply.

We collect and use personal data to provide access to the Springer Nature journal content. We may also use these personal data internally within ResearchGate and Springer Nature and as agreed share it, in an anonymised way, for purposes of tracking, analysis and reporting. We will not otherwise disclose your personal data outside the ResearchGate or the Springer Nature group of companies unless we have your permission as detailed in the Privacy Policy.

While Users may use the Springer Nature journal content for small scale, personal non-commercial use, it is important to note that Users may not:

1. use such content for the purpose of providing other users with access on a regular or large scale basis or as a means to circumvent access control;
2. use such content where to do so would be considered a criminal or statutory offence in any jurisdiction, or gives rise to civil liability, or is otherwise unlawful;
3. falsely or misleadingly imply or suggest endorsement, approval , sponsorship, or association unless explicitly agreed to by Springer Nature in writing;
4. use bots or other automated methods to access the content or redirect messages
5. override any security feature or exclusionary protocol; or
6. share the content in order to create substitute for Springer Nature products or services or a systematic database of Springer Nature journal content.

In line with the restriction against commercial use, Springer Nature does not permit the creation of a product or service that creates revenue, royalties, rent or income from our content or its inclusion as part of a paid for service or for other commercial gain. Springer Nature journal content cannot be used for inter-library loans and librarians may not upload Springer Nature journal content on a large scale into their, or any other, institutional repository.

These terms of use are reviewed regularly and may be amended at any time. Springer Nature is not obligated to publish any information or content on this website and may remove it or features or functionality at our sole discretion, at any time with or without notice. Springer Nature may revoke this licence to you at any time and remove access to any copies of the Springer Nature journal content which have been saved.

To the fullest extent permitted by law, Springer Nature makes no warranties, representations or guarantees to Users, either express or implied with respect to the Springer nature journal content and all parties disclaim and waive any implied warranties or warranties imposed by law, including merchantability or fitness for any particular purpose.

Please note that these rights do not automatically extend to content, data or other material published by Springer Nature that may be licensed from third parties.

If you would like to use or distribute our Springer Nature journal content to a wider audience or on a regular basis or in any other manner not expressly permitted by these Terms, please contact Springer Nature at

onlineservice@springernature.com