

Streamlining analysis methods for large acoustic surveys using automatic detectors with operator validation

Thomas Webber¹  | Douglas Gillespie¹  | Timothy Lewis²  | Jonathan Gordon¹  |
Tararak Ruchirabha³ | Kirsten F. Thompson^{4,5} 

¹Sea Mammal Research Unit, Scottish Oceans Institute, University of St. Andrews, St. Andrews, UK

²Ale Oak Cottage, Ale Oak, UK

³Greenpeace International, Amsterdam, The Netherlands

⁴Biosciences, College of Life & Environmental Sciences, University of Exeter, Exeter, UK

⁵Greenpeace Research Laboratories, University of Exeter, Exeter, UK

Correspondence

Douglas Gillespie

Email: dg50@st-andrews.ac.uk

Handling Editor: Chloe Robinson

Abstract

1. Passive acoustic surveys are becoming increasingly popular as a means of surveying for cetaceans and other marine species. These surveys yield large amounts of data, the analysis of which is time consuming and can account for a substantial proportion of the survey budget. Semi-automatic processes enable the bulk of processing to be conducted automatically while allowing analyst time to be reserved for validating and correcting detections and classifications.
2. Existing modules within the Passive Acoustic Monitoring software PAMGuard were used to process a large (25.4 Terabyte) dataset collected during towed acoustic ship transits. The recently developed 'Multi-Hypothesis Tracking Click Train Detector' and the 'Whistle and Moan Detector' modules were used to identify occasions within the dataset at which vocalising toothed whales (odontocetes) were likely to be acoustically present. These putative detections were then reviewed by an analyst, with false positives being corrected. Target motion analysis provided a perpendicular distance to odontocete click events enabling the estimation of detection functions for both sperm whales and delphinids. Detected whistles were assigned to the lowest taxonomical level possible using the PAMGuard 'Whistle Classifier' module.
3. After an initial tuning process, this semi-automatic method required 91 hr of an analyst's time to manually review both automatic click train and whistle detections from 1,696 hr of survey data. Use of the 'Multi-Hypothesis Tracking Click Train Detector' reduced the amount of data for the analyst to search by 74.5%, while the 'Whistle and Moan Detector' reduced data to search by 85.9%. In total, 443 odontocete groups were detected, of which 55 were from sperm whale groups, six were from beaked whales, two were from porpoise and the remaining 380 were identified to the level of delphinid group. An effective survey strip half width of 3,277 and 699 m was estimated for sperm whales and delphinids respectively.

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2022 The Authors. *Methods in Ecology and Evolution* published by John Wiley & Sons Ltd on behalf of British Ecological Society.

4. The semi-automatic workflow proved successful, reducing the amount of analyst time required to process the data, significantly reducing overall project costs. The workflow presented here makes use of existing modules within PAMGuard, a freely available and open-source software, readily accessible to acoustic analysts.

KEYWORDS

acoustic survey, acoustics, bioacoustic, click train detector, open-source, PAMGuard, semi-automatic

1 | INTRODUCTION

Researchers are increasingly using bioacoustics to monitor remote marine ecosystems. Passive acoustic monitoring (PAM) has been used to study cetaceans for several decades using a range of approaches. PAM methods allow researchers to investigate a variety of ecological questions across a range of taxa including density estimation (Leaper et al., 2000, 2003; Marques et al., 2009), spatial and temporal distributions (e.g. Merkens et al., 2019; Todd et al., 2020), effects of anthropogenic activity and potential mitigation (e.g. Baumgartner et al., 2019; Macaulay et al., 2017; Malinka et al., 2018) as well as supporting visual surveys (Barlow & Taylor, 2005; Gridley et al., 2020). Mobile sampling, using towed arrays (e.g. Gordon et al., 2020; Rone et al., 2014; Thode, 2004) or gliders (e.g. Bittencourt et al., 2018; Cauchy et al., 2020) can provide designed coverage of larger survey areas. Charter and running costs for dedicated vessels can become a major budgetary component for towed array surveys. However, the use of automated data collection systems can allow towed PAM surveys to be carried out from platforms of opportunity with little human supervision of data collection. Platforms of opportunity are vessels at sea for other purposes which can allow ancillary data collection. Typically, vessel costs are already covered by the vessel's primary task, thus platform of opportunity surveys can be very cost effective. Using platforms of opportunity largely eliminates vessel costs but highlights the importance of cost-effective analysis of large passive acoustic datasets where the project budget is now dominated by analyst time rather than survey costs.

Long-term platform of opportunity projects and autonomous static recorders can collect datasets which extend over months or years. This is compounded by the high sample rates required to capture high frequency vocalisations such as the ~130kHz echolocation clicks of harbour porpoise (Clausen et al., 2011; Villadsgaard et al., 2007) and *Cephalorhynchus* dolphins (Kyhne et al., 2009), further increasing the volume of data to process.

In the absence of automated detectors, data are typically processed by an analyst viewing a short-term Fourier transform (STFT) on a spectrogram display and listening to sections of interest. This takes considerable time, particularly for high frequency data where an analyst must listen in less than real time. Semi-automated analysis reduces the manual effort required from analysts, with the analyst's time being reserved for making the final decisions on targeted data.

Automation can also reduce biases and errors which often result with human analysts (Aide et al., 2013; Heinicke et al., 2015) and provide additional information from the data such as bearings calculated from time of arrival differences for signals received on multiple hydrophones. A suite of automatic processes are currently available for analysing acoustic data for a range of species, including click detectors and click classifiers (Gillespie et al., 2008; Madhusudhana et al., 2015; Miller & Miller, 2018), energy band comparisons (Klinck & Mellinger, 2011), extraction of spectral features (Gillespie et al., 2013; Lin & Chou, 2015) and more recently, machine learning methods (Bergler et al., 2019; Bermant et al., 2019; Jiang et al., 2018; Shamir et al., 2014). These methods differ in their computational requirements, performance and ability to process sounds from a range of species, the wider environment and anthropogenic sources.

All automatic detection methods are subject to errors: both false negatives where detections from the species of interest are missed, and false positives where detections are erroneously assigned to the species of interest. Choices about how to balance these errors and their impact on the overall results are dependent on the study. Density estimation methods based on distance sampling (Buckland et al., 1993) which were first developed for visual surveys, generally deal well with missed detections by directly estimating the probability of detection as a function of distance from the track line. As long as there is a high probability of detection along the survey track, missed detections (false negatives) at greater distances are of no consequence since the reduction in detections is measured by the reduction in the estimate of detection probability (González et al., 2018; Thomas & Marques, 2012). Marques et al. (2009), showed that acoustic data can also be used if a false positive rate is known, but this is likely to vary for different datasets based on the characteristics of interfering noise. Therefore, determining whether the analyst should examine all detections and remove false detections, or examine a subset and estimate the fraction of detections that were false positives, will depend on deciding a balance between endeavour and statistical robustness.

Sperm whales lend themselves well to acoustic surveys. Their loud clicks can be heard at distances of several km, and they can be tracked and localised from a moving vessel with very modest equipment using target motion methods. Several studies have published abundance estimates for sperm whales using standard line transect survey approaches (Lewis et al., 1998, 2007, 2018) or methods with

small modifications (Barlow et al., 2001; Barlow & Taylor, 2005). Data are often processed with either RainbowClick (Gillespie & Leaper, 1996) or more recently PAMGuard (Gillespie et al., 2008). Both programs combine a simple transient click detector, combined with a sophisticated user interface, to enable an operator to efficiently select and group click trains on consistent bearings likely to come from an individual or closely associated group. The click detector generally produces many false positive detections, which may come from a variety of sources: other cetacean species, propeller and engine noise from the survey vessel and other craft, and other naturally occurring sounds such as breaking waves. It is therefore necessary for an operator to examine every screen page of data to eliminate false detections. However, by only displaying detections, the page length can be longer and much less cluttered than would be possible with a standard spectrogram, and additional information such as bearings to detected sounds are displayed. This makes it possible for an operator to scan data offline at many times real-time.

In this study, we took a total of 1,696 hr of continuous data collected from a hydrophone array deployed from a platform of opportunity while it made routine passages for other purposes. It is hoped that this opportunistic PAM data collection will be the start of a long-term project to collect PAM data on a wide scale, thereby contributing to world-wide cetacean population monitoring efforts. Data were processed for the extraction of multiple classes of cetacean sounds including sperm whale echolocation clicks, broadband delphinid echolocation clicks, narrow band high frequency (NBHF) echolocation clicks and delphinid whistles. We report on the distribution of these sound types and provide detection functions for sperm whales and delphinids along the survey track. Importantly, we demonstrate how a carefully selected combination of automatic and

manual processing allows for the time efficient processing of large datasets.

2 | MATERIALS AND METHODS

2.1 | Data collection

Acoustic data were collected opportunistically on-board M/V *Arctic Sunrise* during passages in the Atlantic, Southern, Arctic and Indian Oceans using a towed hydrophone array (Vanishing Point Ltd; Figure 1). The data collected during this research was collected under permit number RWS-2019/40813 for all work in Antarctic waters. In other regions, no fieldwork permission was required. The array's streamer section comprised two pairs of hydrophone elements mounted within an oil (Isopar M) filled 5 m long, flexible, 35 mm diameter polyurethane tube. This was towed at the end a 350m Kevlar-strengthened tow cable. Two hydrophones, the 'medium frequency' pair (Benthos AQ4 elements and Magrec HP02 preamplifiers, nominal frequency range 50Hz to 40kHz) were spaced 3 m apart while the 'high frequency' pair (Magrec HP03 hydrophone and preamplifiers units, nominal frequency range 1–200kHz) were separated by 50cm. Each array element was connected to one channel of a four-channel SAIL data acquisition card (St Andrews Instrumentation Ltd) where analogue filtering and gain were applied before each channel was sampled at 500kHz. High pass filters of 10 Hz and gain of 6 dB were applied to the 'medium frequency' channels 0 and 1, while a high pass filter of 2 kHz and gain of 12 dB applied to each of the 'high frequency' channels 2 and 3. Data from the SAIL acquisition card were written as four channel .wav files using PAMGuard (Gillespie et al., 2008) (available at www.pamguard.org),

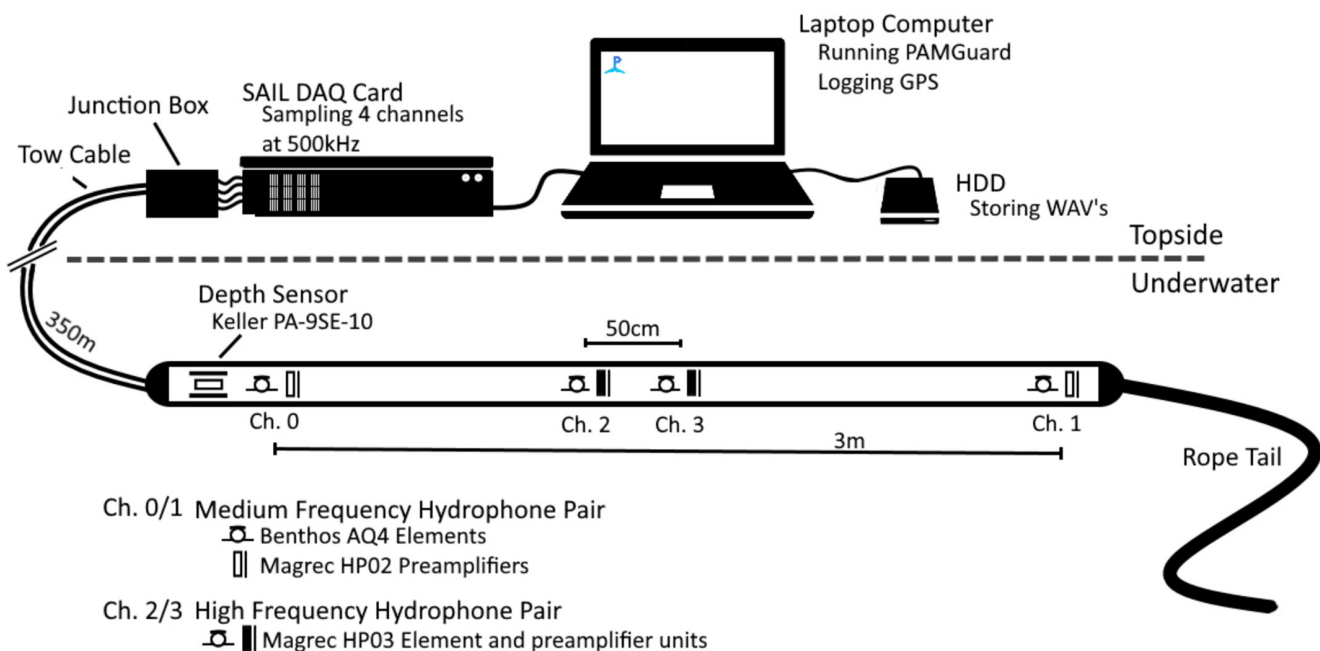


FIGURE 1 Schematic of the towed PAM array and recording system used onboard M/V *Arctic sunrise* during acoustic surveys

which also carried out real-time acoustic processing, displayed results and logged the ship location from GPS.

2.2 | Click processing

The raw wav data files were reprocessed onshore in conjunction with GPS data collected during the survey using PAMGuard (version 2.01.05).

2.2.1 | Detection and detector configuration

Odontocete clicks were detected on recordings from the high frequency hydrophones (channels 2 and 3) using a PAMGuard click detector module. Time of arrival differences for the signal on the two hydrophones were used to estimate an angle of arrival for each detected click relative to the hydrophone array.

To achieve a good compromise between detection efficiency and processing workload, an exploratory analysis was conducted on a representative subset of data from transits 1 and 2, to determine the PAMGuard click detector settings which enabled detection of all manually identified vocalisations while removing as many detections from noise sources as possible. This subset was representative of typical encounters identified during an initial pass through the data. The vessel propulsion system (propeller and engine noise) was the source of many false detections. PAMGuard allows detections on a certain range of bearings to be vetoed. All detections in a 40-degree sector ahead of the vessel, with bearings between +20 and -20 were discarded. To further reduce false detections from background noise, a range of detector trigger thresholds between 10 and 19 dB in 3 dB steps were applied and a value was chosen where the maximum number of false detections were removed, while still retaining all but one odontocete click train. This optimised threshold was used to reanalyse all the recordings and timing, bearing and waveform information for the detected clicks were written to PAMGuard output files for further review and analysis.

Cetacean click vocalisations typically occur in trains, with fairly consistent and characteristic inter click intervals. Thus, trains of clicks, on a consistent bearing, are usually a more reliable cue than individual clicks. The 'Multi-Hypothesis Tracking (MHT) Click Train Detector' module within PAMGuard (Macaulay, 2020) was used to automatically group detected clicks into click trains. The module assesses the Inter-Click-Interval (ICI), amplitude, frequency content and bearing information of clicks to assemble putative clicks trains and then calculates the likelihood of being a true click train for every possible click combination. As more clicks are included in the model, the number of possible click trains increases exponentially and a pruning process is implemented so that only the most likely combination of clicks are retained in putative click trains. Each click train is given a χ^2 score. The lower the score, the more likely it is that the clicks within the train come from the same source or target animal.

This process is computationally intensive, and while the pruning process increases the efficiency of the model (Macaulay, 2020), removing as many false positive clicks as possible before running the MHT click train detector proved essential.

Settings for the MHT click train detector were based on those suggested by Macaulay (2020) and adjusted iteratively for the data subset. The settings for the MHT click train detector and its classifier were then validated against a full manual analysis of the subset using existing MATLAB functions for PAMGuard (available at: <https://github.com/PAMGuard/PAMGuardMatlab>) within custom MATLAB scripts (version 9.9.0, MATLAB, 2020).

2.2.2 | Classification

Following previous work to identify beaked whales in similar acoustic surveys (Keating & Barlow, 2013; Rone et al., 2014; Yack et al., 2010), two narrow band click classifiers with frequency sweeps were applied to detect beaked whales. The first using the PAMGuard defaults for beaked whales with a test band between 24 and 48 kHz, and the second higher frequency test band (40–80 kHz) to search for higher frequency beaked whale clicks. The presence of a frequency sweep, assessed by eye in Wigner plots of individual clicks was useful in identifying beaked whales. A narrow band classifier was also used to detect narrow band high frequency (NBHF) clicks, with a test band between 100 and 150 kHz, providing a classifier for any NBHF species such as harbour porpoise *Phocoena phocoena*, dwarf and pygmy sperm whales (*Kogia* spp.) and NBHF delphinids (e.g. *Cephalorhynchus* spp.). The classifier within the MHT module uses spectral template classifiers which correlate the average spectrum of each click train with species specific spectral templates and inter-click interval parameters. Classifiers were run within the MHT click train detector for sperm whales, beaked whales and dolphins.

2.3 | Whistle processing

PAMGuard's 'Whistle and Moan Detector' (Gillespie et al., 2013) was run to detect odontocete whistle contours up to 24 kHz on the wav data files from the 'medium frequency' hydrophone pair (decimated to 48 kHz), using settings provided in Gillespie et al. (2013). The detector identifies tonal sounds within recordings using a multi-stage process which removes noise, calculates an FFT, applies an amplitude threshold and joins narrow band peaks in FFTs which are close in time and frequency to show 'whistle contours'.

Whistle contours were then classified to the species level using PAMGuard's whistle classifier (Gillespie et al., 2013). The classifier works by breaking up the detected contours into fragments of equal length before the mean frequency (Hz), frequency slope (Hz/s) and curvature (Hz/s²) of the fragment are extracted. The distribution patterns of these parameters are calculated for whistles in encounters. The mean, standard deviation and skew of these parameters have been shown to vary between species (Gillespie et al., 2013).

Thus, whistles can be classified from multiple whistle fragments, by comparing distributions of fragments measured during acoustic encounters with distributions of contours from known species. Whistle classification was run separately for different geographical regions likely to have a different combination of 'whistling' species. In each region the classifier was trained using pre-existing and pre-labelled whistle contours of species likely found in that region. The training data were not necessarily collected from that region however. Contours used in training, also used in Gillespie et al. (2013), had been sampled at 48 kHz with fragment length and section length parameters set at 30 bins (160 ms) and 60 fragments respectively.

Where an event could not be classified, due to an insufficient number of whistle fragments, a range of frequency metrics of each event, such as mean whistle frequency and mean whistle slope were extracted. This information was used to aid species identification. Identification to the family level was attempted where species level identification was not possible.

2.4 | Manual audit

All sections of data that contained click train detections and/or whistle detections were manually audited using the PAMGuard viewer displays and data map. The automatic classification of triggered click trains (e.g. sperm whale, beaked whale, NBHF, delphinid) was corrected where necessary. Echolocating odontocetes can most easily be distinguished on the bearing-time display of the click detector and click characteristics allow events to be placed into a species

group. For example, using the PAMGuard Wigner plot for upsweep verification of beaked whale clicks (Papandreou-Suppappola & Antonelli, 2001; Yack et al., 2013). The MHT click train detector often fragmented a single click train into separate sections. In these cases, the analyst 'marked up' trains more accurately (Figure 2).

As the whistle and moan detector can trigger on any tonal sound within its detection range, it was important to inspect detected contours to ensure only those from delphinids were included in later analyse and labelled using PAMGuard's 'Spectrogram Annotation' module. Delphinid click trains and whistles were merged where temporal overlap occurred into delphinid encounter events.

2.5 | Localisation

Click trains were localised using PAMGuard's Target Motion Analysis (TMA) module using the two-dimensional simplex method. This minimises the least squares error within a click train of bearings to a stationary location, estimating a different location for each side of the track. A simple two element linear array was used in these transits. The bearings calculated by a time of arrival difference actually place the target on a semi-circular arc passing beneath the vessel's track line. Fortunately, line transect surveys are quite forgiving for these ambiguities. It has been shown (Leaper et al., 1992; Lewis et al., 2018) that when the perpendicular distances to detections are typically greater than the likely depth of the detected animal, the 'vertical ambiguity' can be accounted for by the detection function and has little impact on

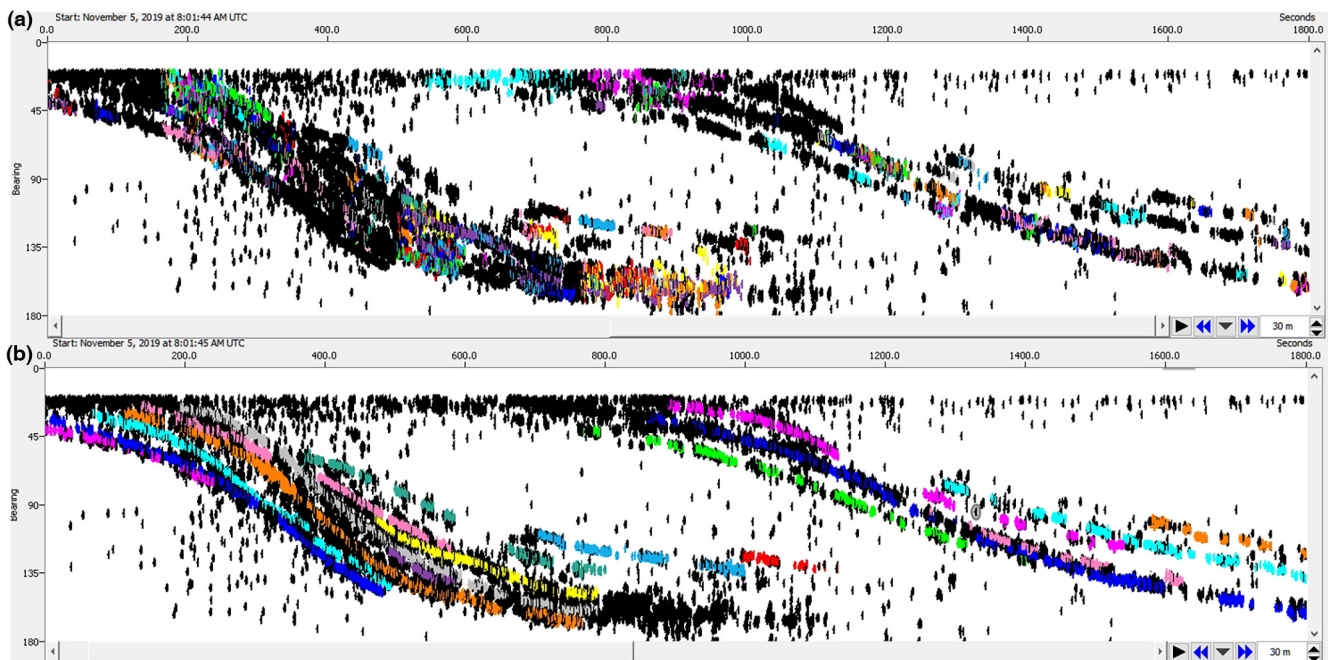


FIGURE 2 Bearing-time windows within PAMGuard showing a sperm whale encounter over a 30-min period. Window A shows the fragmented trains produced by the click train detector, with window B showing the same event after a manual revision and mark-up process to identify single click trains for each vocalising whale where possible. Due to the overlap of click trains, especially at the upper and lower ends of the bearing scale, it is not always possible to distinguish between vocalising individuals

density estimation. Thus, for localisation purposes bearings can be considered as being horizontal. A left right uncertainty still remains, however, distance sampling methodology only requires a perpendicular distance from the track line (Buckland et al., 1993). Thus, there is no requirement to know which side of the track line an animal is on.

Perpendicular distances can be used to determine a detection function, an equation showing how the probability of detection falls with range from the track line. Perpendicular distances to click trains were processed using the 'DISTANCE' package (version 1.0.2, Miller, 2020) within the R statistical analysis environment (version 4.0.2, R Core Team, 2020) to estimate detection functions. Detection functions were calculated for both sperm whales and for non-NBHF delphinids.

Dolphins often bow ride and mill around the vessel. Clearly, target motion analysis cannot be used in these cases. However, a 'delphinid' detection function was calculated using a subset of detections which clearly moved past the array that had well-defined click trains.

Detection functions were calculated using data from all transits, as the same vessel and equipment were used throughout the study. Half normal and hazard rate models were explored for each species group, and the best model for each selected based on Akaike information criterion (AIC) scores.

3 | RESULTS

A total of 1,696 hr of four channel acoustic data were collected during more than 30,000km of survey effort across the Atlantic, Southern, Arctic and Indian Oceans (Table 1). This resulted in 25.4 Terabytes (TB) of 16 bit .wav files.

3.1 | Click processing

After exploratory analysis, a 16 dB click detection threshold was chosen over that of the 10 dB threshold used for real-time processing and other tested thresholds (13 and 19 dB) on the basis of the number of retained odontocete clicks and number of noise-originating clicks removed. The processed binary files produced by PAMGuard using a 16 dB compared to a 10 dB threshold were reduced in size between 78.5% and 99.6% for each transit (Table 2). Processing the 1,696.3hr of .wav files in PAMGuard took approximately 353 hr, an average of 4.8x real-time.

The MHT click train detector found 5,531 click trains within the subset, 636 of which were automatically classified as sperm whales. A spectral template threshold of 0.7 was chosen. This was the largest value which retained all true positives. Eighty-six sperm whale click trains were manually identified in the subset of data. All of these were identified by the automated click train detector, and no manually identified sperm whale click trains were missed. Five hundred and fifty click trains were, therefore, incorrectly

identified as sperm whale click trains. False positive click trains were a mixture of echoes from true sperm whale events, delphinid clicks with peaks at lower frequency and detections from the noise of the towing vessel. 1,144 of the 5,531 click trains from the subset were classified as delphinids, all of these detections were part of manually audited delphinid events and no manually identified delphinid event was missed by the MHT click train detector. Most of the 3,751 click trains which were not classified to species were from the towing vessel on occasions when vessel noise occurred outside the 40° veto zone, such as reflections from the seabed in shallow water and when the vessel was changing heading. Another source of unclassified click trains were short segments of sperm whale click trains which did not meet the correlation threshold of the spectral classifier. In all cases, there were many other short click train segments which did meet the classifier threshold, and so no sperm whale event was missed. Based on this exploratory analysis from the subset, unclassified click trains were discarded from any further analyses and were not manually validated when processing the entire dataset.

The total recording effort during the subset was 52.3 hr, while the duration of the click train encounters (identified automatically and highlighted for the analyst's attention and interpretation) was 17.7 hr, a reduction of approximately 66%, with none of the odontocete detections identified during the initial manual analysis being missed. Analysing this reduced dataset took the analyst 2.2 hr rather than 4.3 hr for the fully manual mark up, almost halving the workload.

The MHT click train detector settings were applied across the entire dataset. A standard laptop computer (Intel i7 2.80GHz 2nd generation processor with 8GB of RAM) running the PAMGuard MHT click train detector unsupervised took 22.3hr to process 1,696.3hr of data. Across the entire dataset the MHT click train detector reduced the amount of data to process manually by 74.5%.

The three click classifiers with a frequency sweep took 56 min to classify all detected clicks for the entire dataset. In total, the manual effort marking up classified clicks from the click classifier and click trains from the click train detector took 55.2hr for the entire dataset.

It is likely, based on the acoustic characteristics of their clicks, that the six beaked whale detections in the Atlantic were of Cuvier's *Ziphius cavirostris*, Blainville's *Mesoplodon densirostris* or True's beaked whales *Mesoplodon mirus* (Baumann-Pickering et al., 2013; DeAngelis et al., 2018; Shaffer et al., 2013). There were 19 NBHF events detected in total. Two of which were likely harbour porpoise *Phocoena phocoena* off the northern Norwegian coast (Clausen et al., 2011; Quintela et al., 2020; Storrie et al., 2018; Villadsgaard et al., 2007). Sixteen detections made in shallow waters along the South American coast and one in the Drake passage were likely species of Cephalorhynchus dolphins. Thought to be Commerson's *Cephalorhynchus commersonii*, Peale's *Lagenorhynchus australis* or hourglass dolphins *Lagenorhynchus cruciger* based on click characteristics and species distribution (Cipriano, 2018; Dellabianca et al., 2012; Kyhn et al., 2009, 2010; Reyes Reyes et al., 2015). From

TABLE 1 Summary of acoustic data collection on board M/V *Arctic sunrise* by transit with data size given in gigabytes (GB)

Transit	Start-End	Recording effort (hours)	Size of recordings (GB)	Number of WAV files	Distance covered (km)	Average speed (knots)
1) English Channel to Islas Canarias	22-Aug-2019 28-Aug-2019	137	1,971	1,874	2,659	8.1
2) Islas Canarias to Dakar	01-Sep-2019 04-Sep-2019	52	753	1,051	1,264	8.9
3) Dakar to Cape Town	16-Sep-2019 15-Oct-2019	191	2,732	3,661	4,057	5.9
4) Cape Town to Vema to Cape Town	23-Oct-2019 07-Nov-2019	79	2,133	1,598	3,903	7.6
5) Cape Town to South Atlantic	03-Dec-2019 16-Dec-2019	261	3,763	5,025	4,783	10.4
6) South Atlantic to Ushuaia	17-Dec-2019 26-Dec-2019	218	3,143	4,827	4,783	10.4
7) Ushuaia to WAP	06-Jan-2020 10-Feb-2020	180	2,596	3,338	3,138	8.6
8) Norway to Svalbard	03-Sep-2020 30-Sep-2020	91	1,317	3,039	2,114	9.5
9) Seychelles to Saya de Malha	02-Mar-2021 30-Mar-2021	487	7,006	10,223	4,961	7.4
Total		1,696	25,414	34,636	31,662	8.5

TABLE 2 Number of detected clicks in millions and their file size in gigabytes (GB) to highlight reduction in noise when a 16 dB click detector threshold and 20° veto was applied over the default 10 dB threshold

Transit	Number of detected clicks in millions (GB)		% reduction
	10 dB	16 dB	
1	62.0 (13.0)	2.2 (0.5)	96.4
2	23.7 (5.0)	0.6 (0.1)	97.6
3	93.2 (21.7)	0.7 (0.2)	99.2
4	76.2 (19.8)	0.3 (0.1)	99.6
5	14.4 (3.0)	0.2 (0.03)	98.9
6	14.9 (3.1)	3.2 (0.7)	78.5
7	18.7 (3.9)	1.1 (0.2)	94.3
8	8.0 (1.7)	2.7 (0.1)	97.1
9	56.2 (11.8)	5.1 (1.1)	90.9
Total	367.3 (83.0)	16.1 (3.0)	95.6

these 16 acoustic encounters, two were confidently attributed to Peale's dolphins on the basis of sightings made during the acoustic encounter.

3.2 | Whistle processing

PAMGuard's whistle classifier took approximately 62 min to process the entire dataset. The mean correct classification rate in the training

dataset varied for each regional species group (64.1%–82.9%) due to the different mix of species included for each region. Confusion matrices showed that within the training dataset all species were correctly classified on most occasions. However, some species showed consistently higher (>20%) miss-classification rates across regional groups. For example, pilot whales were miss-classified as killer whales between 29.7% and 32.5% of the time.

Two hundred and one of the 349 manually verified whistle events detected during this project were classified by the whistle classifier, with most of those not classified containing fewer than 10 whistle contours. For example, four delphinid detections in the Western Antarctic Peninsula (WAP) were not classified but were manually attributed to killer whales *Orcinus orca* based on vocalisation characteristics (Reyes Reyes et al., 2017; Trickey et al., 2014; Wellard et al., 2020).

Manual validation of the automatically detected whistles across all transits took approximately 35.7 hr with the whistle and moan detector reducing the amount of data to search by 86%.

Table 3 summarises all acoustic events resulting from automatic click and whistle detection. Classification was based on a combination of sightings data, aural inspection and automated click and whistle classifier outputs.

The best fitting detection functions for both sperm whales and non-NBHF delphinids used a hazard rate model with two parameters and no adjustments. Fit was assessed using minimum AIC. This gave an effective strip half-width of 3,277 m for sperm whales and 699 m non-NBHF delphinids (Table 4). For sperm whales, there is a reduction in the number of detections within ~500 m of the track line (Figure 3).

Species grouping		Common name	Number of acoustic group events
<i>Physeteridae</i>	<i>Physeter macrocephalus</i>	Sperm whale	55
<i>Ziphiidae</i>	Unclassified	Beaked whale	6
<i>Phocoenidae</i>	<i>Phocoena phocoena</i>	Harbour porpoise	2
<i>Delphinidae</i>	Of which:		380
	<i>Delphinus delphis</i>	Common dolphin	8
	<i>Feresa attenuata</i>	Pygmy killer whale	18
	<i>Globicephala</i> spp.	Pilot whale	3
	<i>Grampus griseus</i>	Risso's dolphin	62
	<i>Lagenorhynchus acutus</i> or <i>L. albirostris</i>	White beaked or white sided dolphins	7
	<i>Orcinus orca</i>	Killer whale	6
	<i>Pseudorca crassidens</i>	False killer whale	4
	<i>Stenella coeruleoalba</i>	Striped dolphin	7
	<i>Stenella frontalis</i>	Atlantic spotted dolphin	13
	<i>Stenella longirostris</i>	Spinner dolphin	66
<i>Tursiops truncatus</i>	Bottlenose dolphin	5	
NBHF	<i>Lagenorhynchus australis</i> , <i>L. cruciger</i> or <i>Cephalorhynchus commersonii</i>	Peale's, Hourglass or Commerson's dolphin	17
<i>Delphinidae</i>	Of which: <i>L. australis</i>	Peale's dolphin	2
	Unclassified		164
Total			443

TABLE 3 Summary of acoustic detections made from M/V *Arctic sunrise*. Likely classifications were made using a combination of sightings data, aural inspection and automated whistle and click classifier outputs

Model	AIC score	Effective strip half width (m)	CV of estimate	95% CI of ESHW
Sperm whale				
Hazard rate (hr)	4,834	3,277	0.06	3,135–3,419
Half-normal (hn)	4,852	2,868	0.04	2,456–3,010
Non-NBHF delphinids				
Hazard rate (hr)	1,849	699	0.13	569–829
Half-normal (hn)	1,889	1,222	0.04	1,759–1,352

TABLE 4 Summary of detection functions for sperm whales and non-NBHF delphinids across the entire M/V *Arctic sunrise* acoustic transits using click detections

4 | DISCUSSION

The semi-automatic processes applied to the large acoustic dataset in this study used existing modules within PAMGuard. This methodology significantly improved acoustic analyst efficiency. Semi-automated analysis methods, such as these, are essential for large-scale acoustic surveys of wild animal populations, where data may be gathered near continuously over months and can total tens of terabytes, and budgetary limitations mean that specialist analyst time must be used cost-efficiently.

We appreciate that the analyses of such data will never be completely optimised for a number of reasons. All detection systems suffer from false positive and false negative detections. While a

false positive is easily defined, the definition of a false negative in an acoustic survey is much more complicated. Lowering detection thresholds in order to detect more distant sounds will inevitably lead to a high rate of false positives. False negatives will inevitably increase as detector thresholds are increased to remove false positives. However, distance sampling methodology, used in this study, generally deals well with false negatives (Thomas & Marques, 2012). The detection threshold chosen for the click detector was appropriate for the somewhat noisy towing vessel. Although false positives can generally be removed through manual audit, the additional cost of processing more detections can outweigh the statistical benefit of a larger sample size. A simple 'optimal' configuration is therefore illusory unless we know what we are optimising for: is it the maximum

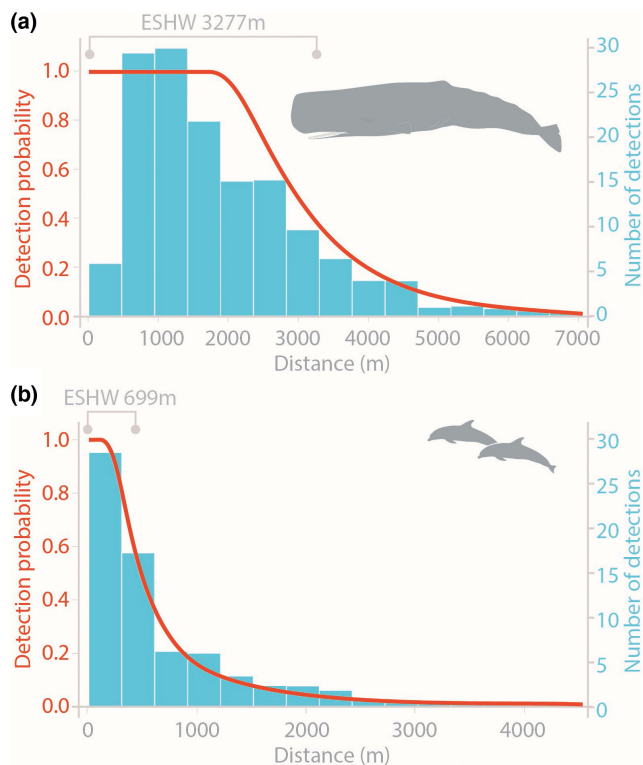


FIGURE 3 Combined hazard rate detection functions for sperm whales and non-NBHF delphinids across all transits on the M/V Arctic sunrise. With effective strip half-width of 3,277 m for sperm whales and 699 m for non-NBHF delphinids

number of detections, the maximum ratio between the numbers of true detections and false positives, the maximum number of detections per unit of overall monetary cost, or what? Scientifically, it is important to have an adequate sample for statistical analysis, with few false positives, requiring analyst input. A more obtainable goal is to achieve the best balance between false positives and false negatives, using appropriate threshold levels which can provide sufficient true detections for statistical analysis while managing the analyst time required to eliminate false positives. If these criteria can be met with reduced human effort, then a 'sweet spot' will be achieved between true detections and false positives that statistically describes animals along the track line.

It is likely that a mixture of automated processing and human validation will remain the standard method for acoustic data analysis for the foreseeable future. However, improved algorithms, with a higher efficiency and/or a lower false alarm rate will reduce the amount of human effort required for data analysis. For example, Shiu et al. (2020) showed that a new algorithm for detecting the calls of North Atlantic right whales, reduces the false alarm rate to the point where a human analyst can check 1 month of continuous data in 7.44 hr. However, potential short comings of automated detections must be kept in mind, for example the potential to miss changes in vocalisations over time (Sirvic, 2015), and for poorly understood or rare species to be missed (Shiu et al., 2020).

The MHT click train detector produced a large number of false positive sperm whale click trains in the subset, but did not miss any

sperm whale events, ensuring important data were not lost during the analysis. Detector settings were heavily influenced by the background noise from the towing vessel and so thresholds should be adjusted to suit the background noise within each particular study.

Validation against the fully manual analysis of the subset showed that all automatically detected delphinid click trains occurred within manually marked up delphinid events and no delphinid events were missed. These comparisons provided a high degree of confidence that few delphinid events would be missed using the MHT click train detector. This detector provides a streamlined way to detect delphinids in large datasets. Grouping individual clicks into click trains and measuring them improves classification and with continued development, there is potential to more accurately associate clicks to individuals (Macaulay, 2020).

While the MHT click train detector does not mark up each click train perfectly, by highlighting click trains automatically, the amount of data an analyst needs to search and mark up is greatly reduced. The 66% reduction in data to search, resulted in a 49% reduction in analyst time. Without reliable detectors, analysis effort is largely governed by data volume; our study shows that with them analysis effort becomes a function of the number of cetacean encounters.

The Whistle and Moan detector (Gillespie et al., 2013) has been used in previous studies to find delphinid whistles (e.g. Erbs et al., 2017; Keating et al., 2015) and vocalisations of baleen whales (e.g. Miller et al., 2016). Gillespie et al. (2013) reported a recall of 88% (the percentage of manual detections which were detected automatically). In this study, false positive whistle detections were manually removed by the analyst. This proved to be an important and necessary step as depth sounders of passing ships and other high frequency signals, likely from telemetry transceivers, triggered the detector on 25 separate occasions. By using this detector, and only investigating triggered events, the amount of data to search was reduced by 85.9% (the analyst needed 35.7 hr to process whistle detections in the full 1,696.3-hr dataset).

For sperm whales, reduction in detection range due to prevailing noise or propagation conditions is shown by the detection function and accounted for in density estimation. For whistles, where we are unable to use target motion analysis to estimate a detection function, it will be necessary to develop alternative methods to measure detection probability.

Although the whistle classifier has been shown to perform well, with Gillespie et al. (2013) reporting correct classification rates up to 94.5%, it is likely some classifications in this study could be incorrect. There is a paucity of specific acoustic data for some species within the survey regions; thus, the data used to train the classifier likely may not have included all species present, nor reflect likely differences in whistles between regions within the same species (Erbs et al., 2017). To increase the confidence in species classification, region specific recordings of as many cetacean species as possible are required to retrain the classifier. This will be particularly important for localised studies estimating abundance or investigating habitat preference (Erbs et al., 2017). When training data become available,

delphinid events within this study can be reclassified, compared and corrected where appropriate.

The combined hazard rate detection function for sperm whales gave an effective strip half-width (ESHW) of 3,277 m. Other studies using similar methodologies have reported ESHW's of 4.2–10 km (Fais et al., 2015; Gordon et al., 2020; Lewis et al., 2018). The 3.3 km ESHW reported is narrower than reported by other studies which is likely due higher levels of noise emitted by the towing vessel in this study. This detection function shows that a vocally active sperm whale within 1,500 m of the track line will be detected, and so $g(0)$ for a vocal sperm whale is equal to 1. However, sperm whales are known to have silent periods which can be a function of social behaviour and may vary regionally (Jaquet et al., 2001; Whitehead & Weilgart, 1991) but these data are limited.

The sperm whale detection function measured during this study showed relatively low detections immediately adjacent to the track line. This is a common characteristic of sperm whale acoustic detection functions derived using data from simple two element towed arrays Gordon et al. (2020). It has been suggested that this is consequence of plotting in two dimensions, which uses bearings that have an unknown vertical component. This effect will be most evident for sperm whales vocalising at depth close to the track line. TMA calculates the distance to the animal from the hydrophone but a substantial component of this will be due to the animal's depth rather than its horizontal distance from the track line. Leaper et al. (1992) and Lewis et al. (2018) explored this with simple simulations and concluded that if the 'shoulders' of the detection function were greater than typical vocalisation depths for the target species then this effect should not lead to substantial biases. Thus, problems could arise if detection range was substantially reduced as a result of high noise levels or unfavourable propagation conditions. If there is a reason to believe that propagation conditions are markedly different in different regions, then efforts should be made to stratify detection functions by region.

There is rather little information on detection range for delphinids during acoustic surveys. Martin et al. (2020) measured an ESHW of 367 m for dusky dolphins *Lagenorhynchus obscurus*, while Rankin et al. (2008) reported that the majority of dolphin detections were within a few kilometres of the array. Given the varied nature of group size and behaviour for delphinid species, and uncertainty in classifications, it is unlikely that reliable estimates of density could be obtained for delphinids using this acoustic method alone.

After an initial tuning process, automatic processing took 268 hr to run unsupervised on standard PC hardware (Intel i7 2.80GHz second generation processor with 8GB of RAM). 90.9 hr of analyst effort was then required to manually audit automatic click train and whistle detections from 1,696 hr of data. A major advantage of PAM surveys on platforms of opportunity is that low field costs can allow very large datasets to be acquired. Analysis costs then become a large proportion of the overall budget. Thus, reducing required analyst time will have a large impact on overall costs allowing such surveys to become a routine activity during transits producing valuable

data on poorly known oceanic species. We suggest a similar approach is applicable for other applications where large bioacoustics datasets are one of the primary survey tools, for example, wild bats, insects or even in domestic animal welfare monitoring (McLoughlin et al., 2019; Zilli et al., 2014).

The coverage provided by this initial survey is extremely broad, but sparse. The real value of initiatives like this will come once data have been collected for several years. Even so, interesting information on distributions in rarely surveyed areas is evident. Delphinids were detected frequently across every transit. Sperm whales were detected close to the shelf break and in oceanic waters in all transits except that in the Southern Ocean. Higher detection rates were evident off north-west Africa, South Africa, South America and Svalbard. Sperm whales were also detected near seamounts such as Vema and Filippov in the South Atlantic and the Bathymetrists Seamounts Chain in the Tropical Atlantic. These detections address key knowledge gaps on the distribution of sperm whales in poorly surveyed ocean areas, helping researchers target future survey efforts. Further survey effort would provide data which could be used in habitat models to enable a more robust comparison between regions and expand on our understanding of fine scale distribution and their drivers. The data gathered during this study are freely available through the online repository OBIS-Seamaps (<https://seamap.env.duke.edu/>).

Over the past decade, more reliable and cost-effective hardware, and sophisticated software has enabled non-specialist researchers to conduct bioacoustic surveys. In the marine environment such acoustic surveys can be conducted during opportunistic transits using a variety of survey platforms using highly automated and relatively inexpensive towed hydrophone systems. The task of detecting and classifying detections in such data so that species distributions and densities can be inferred is a time-consuming process for specialist acoustic analysts. Our study provides a template for efficient analysis of such large-scale acoustic datasets, reducing the time required by specialist analysts, and ultimately the cost of any acoustic-based study.

AUTHORS' CONTRIBUTIONS

K.F.T., D.G. and J.G. conceived the ideas and designed methodology; T.L., K.F.T. and T.R. collected the data; T.W. and D.G. analysed the data; T.W. and D.G. led the writing of the manuscript. All authors contributed to the drafts and gave final approval for publication.

ACKNOWLEDGEMENTS

We thank the crew of the Arctic Sunrise for helping to collect these data during approximately 30,000 km of transits and campaign work in four oceans. In particular, we acknowledge the teamwork involved in the deployment, checking and retrieval of the hydrophone. We also thank José Antonio Vázquez Bonales and Kike Perez Gil for data collection (Amsterdam–Senegal). Logistical support was provided by Grant Oakes, David Santillo, and Paul Johnston along with other members of the Greenpeace Research Laboratories, University of Exeter and Greenpeace Operations Department, Greenpeace International. We also thank

the anonymous reviewers for their comments which helped to improve the clarity of the manuscript.

CONFLICT OF INTEREST

The authors declare that they have no competing interests.

PEER REVIEW

The peer review history for this article is available at <https://publons.com/publon/10.1111/2041-210X.13907>.

DATA AND SOFTWARE AVAILABILITY STATEMENT

Detection data are available from the OBIS-SEAMAP repository <https://seamap.env.duke.edu/dataset/2155> & <https://seamap.env.duke.edu/dataset/2156> (Webber et al., 2021a, 2021b). The PAMGuard software is freely available as an executable program which can be downloaded from www.pamguard.org. Source code is available at <https://github.com/PAMGuard/PAMGuard>. The PAMGuard settings file used in this study can be found at [10.5281/zenodo.6558962](https://doi.org/10.5281/zenodo.6558962) (Webber et al., 2022).

ORCID

Thomas Webber  <https://orcid.org/0000-0003-1367-3688>

Douglas Gillespie  <https://orcid.org/0000-0001-9628-157X>

Timothy Lewis  <https://orcid.org/0000-0001-9369-6948>

Jonathan Gordon  <https://orcid.org/0000-0002-5474-9899>

Kirsten F. Thompson  <https://orcid.org/0000-0003-4277-3549>

REFERENCES

- Aide, T. M., Corrada-Bravo, C., Campos-Cerqueira, M., Milan, C., Vega, G., & Alvarez, R. (2013). Real-time bioacoustics monitoring and automated species identification. *PeerJ*, 2013(1), e103. <https://doi.org/10.7717/peerj.103>
- Barlow, J., Gerrodette, T., & Forcada, J. (2001). Factors affecting perpendicular sighting distances on shipboard line-transect surveys for cetaceans. *Journal of Cetacean Research and Management*, 3(2), 201–212.
- Barlow, J., & Taylor, B. L. (2005). Estimates of sperm whale abundance in the northeastern temperate Pacific from a combined acoustic and visual survey. *Marine Mammal Science*, 21(3), 429–445. <https://doi.org/10.1111/j.1748-7692.2005.tb01242.x>
- Baumann-Pickering, S., McDonald, M. A., Simonis, A. E., Solsona Berga, A., Merckens, K. P. B., Oleson, E. M., Roch, M. A., Wiggins, S. M., Rankin, S., Yack, T. M., & Hildebrand, J. A. (2013). Species-specific beaked whale echolocation signals. *The Journal of the Acoustical Society of America*, 134(3), 2293–2301. <https://doi.org/10.1121/1.4817832>
- Baumgartner, M. F., Bonnell, J., Van Parijs, S. M., Corkeron, P. J., Hotchkiss, C., Ball, K., Pelletier, L., Partan, J., Peters, D., Kemp, J., Pietro, J., Newhall, K., Stokes, A., Cole, T. V. N., Quintana, E., & Kraus, S. D. (2019). Persistent near real-time passive acoustic monitoring for baleen whales from a moored buoy: System description and evaluation. *Methods in Ecology and Evolution*, 10(9), 1476–1489. <https://doi.org/10.1111/2041-210X.13244>
- Bergler, C., Schröter, H., Cheng, R. X., Barth, V., Weber, M., Nöth, E., Hofer, H., & Maier, A. (2019). ORCA-SPOT: An automatic killer whale sound detection toolkit using deep learning. *Scientific Reports*, 9(1), 10997. <https://doi.org/10.1038/s41598-019-47335-w>
- Bermant, P. C., Bronstein, M. M., Wood, R. J., Gero, S., & Gruber, D. F. (2019). Deep machine learning techniques for the detection and classification of sperm whale bioacoustics. *Scientific Reports*, 9(1), 1–10. <https://doi.org/10.1038/s41598-019-48909-4>
- Bittencourt, L., Soares-Filho, W., de Lima, I. M. S., Pai, S., Lailson-Brito, J., Barreira, L. M., Azevedo, A. F., & Guerra, L. A. A. (2018). Mapping cetacean sounds using a passive acoustic monitoring system towed by an autonomous wave glider in the southwestern Atlantic Ocean. *Deep-Sea Research Part I: Oceanographic Research Papers*, 142, 58–68. <https://doi.org/10.1016/j.dsr.2018.10.006>
- Buckland, S. T., Anderson, D. R., Burnham, K. P., & Laake, J. L. (1993). *Distance sampling, estimating abundance of biological populations*. Chapman & Hall.
- Cauchy, P., Heywood, K. J., Risch, D., Merchant, N. D., Queste, B. Y., & Testor, P. (2020). Sperm whale presence observed using passive acoustic monitoring from gliders of opportunity. *Endangered Species Research*, 42, 133–149.
- Cipriano, F. (2018). Peale's dolphin. In B. Würsig, J. G. M. Thewissen, & K. M. Kovacs (Eds.), *Encyclopedia of marine mammals* (pp. 698–701). Elsevier. <https://doi.org/10.1016/b978-0-12-804327-1.00193-x>
- Clausen, K. T., Wahlberg, M., Beedholm, K., Deruiter, S., & Madsen, P. T. (2011). Click communication in harbour porpoises *Phocoena phocoena*. *Bioacoustics*, 20(1), 1–28. <https://doi.org/10.1080/09524622.2011.9753630>
- DeAngelis, A. I., Stanistreet, J. E., Baumann-Pickering, S., & Cholewiak, D. M. (2018). A description of echolocation clicks recorded in the presence of True's beaked whale (*Mesoplodon mirus*). *The Journal of the Acoustical Society of America*, 144(5), 2691–2700. <https://doi.org/10.1121/1.5067379>
- Dellabianca, N., Scioscia, G., Schiavini, A., & Raya Rey, A. (2012). Occurrence of hourglass dolphin (*Lagenorhynchus cruciger*) and habitat characteristics along the Patagonian shelf and the Atlantic Ocean sector of the Southern Ocean. *Polar Biology*, 35(12), 1921–1927. <https://doi.org/10.1007/s00300-012-1217-0>
- Erbs, F., Elwen, S. H., & Gridley, T. (2017). Automatic classification of whistles from coastal dolphins of the southern African subregion. *The Journal of the Acoustical Society of America*, 141(4), 2489–2500. <https://doi.org/10.1121/1.4978000>
- Fais, A., Aguilar Soto, N., Johnson, M., Pérez-González, C., Miller, P. J. O., & Madsen, P. T. (2015). Sperm whale echolocation behaviour reveals a directed, prior-based search strategy informed by prey distribution. *Behavioral Ecology and Sociobiology*, 69(4), 663–674. <https://doi.org/10.1007/s00265-015-1877-1>
- Gillespie, D., Caillat, M., Gordon, J., & White, P. (2013). Automatic detection and classification of odontocete whistles. *The Journal of the Acoustical Society of America*, 134(3), 2427–2437. <https://doi.org/10.1121/1.4816555>
- Gillespie, D., & Leaper, R. (1996). Detection of sperm whale *Physeter macrocephalus* clicks and discrimination of individual vocalizations. *European Research on Cetaceans*, 10, 87–91.
- Gillespie, D., Mellinger, D., Gordon, J., McLaren, D., Redmond, P., McHugh, R., Trinder, P., Deng, X., & Thode, A. (2008). PAMGuard: Semiautomated, open source software for real-time acoustic detection and localisation of cetaceans. *Proceedings of the Institute of Acoustics*, 30(5). <http://toc.proceedings.com/04143webtoc.pdf>
- González, E. S., Camp, R. J., Tanimoto, A. M., De Oliveira, P. M., Lima, B. B., Marques, T. A., & Hart, P. J. (2018). Density estimation of sound-producing terrestrial animals using single automatic acoustic recorders and distance sampling. *Avian Conservation and Ecology*, 13(2), 7. <https://doi.org/10.5751/ACE-01224-130207>
- Gordon, J., Gillespie, D., Leaper, R., Lee, A., Porter, L., O'Brien, J., Meade, R., Á' Cadhla, O., & Berrow, S. (2020). A first acoustic density estimate for sperm whales in Irish offshore waters. *Journal of*

- Cetacean Research and Management, 21(1), 123–133. <https://doi.org/10.47536/jcrm.v21i1.187>
- Gridley, T., Martin, M. J., Slater, J., Roux, J. P., Swift, R. J., & Elwen, S. H. (2020). Towed passive acoustic monitoring complements visual survey methods for Heaviside's dolphins *Cephalorhynchus heavisidii* in the Namibian islands marine protected area. *African Journal of Marine Science*, 42(4), 495–506. <https://doi.org/10.2989/1814232X.2020.1848925>
- Heinicke, S., Kalan, A. K., Wagner, O. J. J., Mundry, R., Lukashevich, H., & Kühl, H. S. (2015). Assessing the performance of a semi-automated acoustic monitoring system for primates. *Methods in Ecology and Evolution*, 6(7), 753–763. <https://doi.org/10.1111/2041-210X.12384>
- Jaquet, N., Dawson, S., & Douglas, L. (2001). Vocal behavior of male sperm whales: Why do they click? *The Journal of the Acoustical Society of America*, 109(5), 2254–2259. <https://doi.org/10.1121/1.1360718>
- Jiang, J., Bu, L., Wang, X., Li, C., Sun, Z., Yan, H., Hua, B., Duan, F., & Yang, J. (2018). Clicks classification of sperm whale and long-finned pilot whale based on continuous wavelet transform and artificial neural network. *Applied Acoustics*, 141, 26–34.
- Keating, J. L., & Barlow, J. (2013). NOAA technical memorandum NMFS summary of PAMGuard beaked whale click detectors and classifiers used during the 2012 Southern California behavioral response study. Retrieved from <https://repository.library.noaa.gov/view/noaa/4535>
- Keating, J. L., Oswald, J. N., Rankin, S. & Barlow, J. (2015). NOAA technical memorandum NMFS whistle classification in the California current: A Complete Whistle Classifier For A Large Geographic Region With High Species Diversity. <https://doi.org/10.7289/V5/TM-SWFSC-552>
- Klinck, H., & Mellinger, D. (2011). The energy ratio mapping algorithm: A tool to improve the energy-based detection of odontocete echolocation clicks. *The Journal of the Acoustical Society of America*, 129(4), 1807–1812. <https://doi.org/10.1121/1.3531924>
- Kyhn, L. A., Jensen, F. H., Beedholm, K., Tougaard, J., Hansen, M., & Madsen, P. T. (2010). Echolocation in sympatric peale's dolphins (*Lagenorhynchus australis*) and commerson's dolphins (*Cephalorhynchus commersonii*) producing narrow-band high-frequency clicks. *Journal of Experimental Biology*, 213(11), 1940–1949. <https://doi.org/10.1242/jeb.042440>
- Kyhn, L. A., Tougaard, J., Jensen, F., Wahlberg, M., Stone, G., Yoshinaga, A., Beedholm, K., & Madsen, P. T. (2009). Feeding at a high pitch: Source parameters of narrow band, high-frequency clicks from echolocating off-shore hourglass dolphins and coastal Hector's dolphins. *The Journal of the Acoustical Society of America*, 125(3), 1783–1791. <https://doi.org/10.1121/1.3075600>
- Leaper, R., Chappell, O., & Gordon, J. (1992). *The development of practical techniques for surveying sperm whale populations acoustically*. Report of the International Whaling Commission, 42.
- Leaper, R., Gillespie, D., Gordon, J., & Matthews, J. (2003). *Abundance assessment of sperm whales*. Paper SC/55/O13 presented to the IWC Scientific Committee, May, 2003, Berlin (unpublished).
- Leaper, R., Gillespie, D., & Papastavrou, V. (2000). Results of passive acoustic surveys for odontocetes in the Southern Ocean. *Journal of Cetacean Research and Management*, 2(3), 187–196.
- Lewis, T., Boisseau, O., Danbolt, M., Gillespie, D., Lacey, C., Leaper, R., Matthews, J., McLanaghan, R., & Moscrop, A. (2018). Abundance estimates for sperm whales in the Mediterranean Sea from acoustic line-transect surveys. *Journal of Cetacean Research and Management*, 18, 103–117.
- Lewis, T., Gillespie, D., Lacey, C., Matthews, J., Danbolt, M., Leaper, R., McLanaghan, R., & Moscrop, A. (2007). Sperm whale abundance estimates from acoustic surveys of the Ionian Sea and straits of Sicily in 2003. *Journal of the Marine Biological Association of the United Kingdom*, 87(1), 353–357. <https://doi.org/10.1017/S0025315407054896>
- Lewis, T., Swift, R., Gonzalbes, P., Bulter, J., & Gordon, J. (1998). *Passive acoustic monitoring of cetacean distribution north-west of the Hebrides*. Retrieved from <https://journal.iwc.int/index.php/jcrm/article/view/187/97>
- Lin, T.-H., & Chou, L.-S. (2015). Automatic classification of delphinids based on the representative frequencies of whistles. *The Journal of the Acoustical Society of America*, 138(2), 1003–1011. <https://doi.org/10.1121/1.4927695>
- Macaulay, J. (2020). *Chapter 2: Open source click train detector for toothed whales*. University of St Andrews.
- Macaulay, J., Gordon, J., Gillespie, D., Malinka, C., & Northridge, S. (2017). Passive acoustic methods for fine-scale tracking of harbour porpoises in tidal rapids. *The Journal of the Acoustical Society of America*, 141(2), 1120–1132. <https://doi.org/10.1121/1.4976077>
- Madhusudhana, S., Gavrilov, A., & Erbe, C. (2015). Automatic detection of echolocation clicks based on a Gabor model of their waveform. *Journal of the Acoustical Society of America*, 137(6), 3077–3086. <https://doi.org/10.1121/1.4921609>
- Malinka, C. E., Gillespie, D. M., Macaulay, J. D. J., Joy, R., & Sparling, C. E. (2018). First in situ passive acoustic monitoring for marine mammals during operation of a tidal turbine in Ramsey sound, Wales. *Marine Ecology Progress Series*, 590, 247–266. <https://doi.org/10.3354/meps12467>
- Marques, T. A., Thomas, L., Ward, J., DiMarzio, N., & Tyack, P. L. (2009). Estimating cetacean population density using fixed passive acoustic sensors: An example with Blainville's beaked whales. *The Journal of the Acoustical Society of America*, 125(4), 1982–1994. <https://doi.org/10.1121/1.3089590>
- Martin, M. J., Gridley, T., Roux, J.-P., & Elwen, S. H. (2020). First abundance estimates of Heaviside's (*Cephalorhynchus heavisidii*) and dusky (*Lagenorhynchus obscurus*) dolphins off Namibia using a novel visual and acoustic line transect survey. *Frontiers in Marine Science*, 7, 800. <https://doi.org/10.3389/fmars.2020.555659>
- MATLAB. (2020). R2020a, version 9.9.0. The MathWorks Inc.
- McLoughlin, M. P., Stewart, R., & McElligott, A. G. (2019). Automated bioacoustics: Methods in ecology and conservation and their potential for animal welfare monitoring. *Journal of the Royal Society Interface*, 16(155), 20190225. <https://doi.org/10.1098/rsif.2019.0225>
- Merkens, K., Simonis, A., & Oleson, E. (2019). Geographic and temporal patterns in the acoustic detection of sperm whales *Physeter macrocephalus* in the central and western North Pacific Ocean. *Endangered Species Research*, 39, 115–133. <https://doi.org/10.3354/esr00960>
- Miller, B. S., Calderan, S., Gillespie, D., Weatherup, G., Leaper, R., Collins, K., & Double, M. C. (2016). Software for real-time localization of baleen whale calls using directional sonobuoys: A case study on Antarctic blue whales. *The Journal of the Acoustical Society of America*, 139(3), EL83–EL89. <https://doi.org/10.1121/1.4943627>
- Miller, B. S., & Miller, E. J. (2018). The seasonal occupancy and diel behaviour of Antarctic sperm whales revealed by acoustic monitoring. *Scientific Reports*, 8(1), 1–12. <https://doi.org/10.1038/s41598-018-23752-1>
- Miller, D. (2020). *Package "distance" (1.0.2)*. Retrieved from <https://github.com/DistanceDevelopment/Distance/>
- Papandreou-Suppappola, A., & Antonelli, L. T. (2001). *Use of quadratic time-frequency representations to analyze cetacean mammal sounds*. Retrieved from <https://apps.dtic.mil/sti/citations/ADA400072>
- Quintela, M., Besnier, F., Seliussen, B., Glover, K. A., & Lindstrøm, U. (2020). Population structure of bycaught harbour porpoise (*Phocoena phocoena*) in Norway. *Marine Biology Research*, 16(3), 141–147. <https://doi.org/10.1080/17451000.2020.1729992>
- R Core Team. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. Retrieved from <https://www.r-project.org/>
- Rankin, S., Barlow, J., Oswald, J., & Ballance, L. (2008). *Acoustic studies of marine mammals during seven years of combined visual and acoustic line-transect*

- surveys for cetaceans in the eastern and Central Pacific Ocean. Technical Memorandum No. NOAA-TM-NMFS-SWFSC-429 (Southwest).
- Reyes Reyes, M., Iñiguez, M., Hevia, M., Hildebrand, J., & Melcon, M. (2015). Description and clustering of echolocation signals of Commerson's dolphins (*Cephalorhynchus commersonii*) in Bahía san Julián, Argentina. *The Journal of the Acoustical Society of America*, 138(4), 2046–2053. <https://doi.org/10.1121/1.4929899>
- Reyes Reyes, M. V., Baumann-Pickering, S., Simonis, A., Melcón, M. L., Trickey, J., Hildebrand, J., & Iñiguez, M. (2017). High-frequency modulated signals recorded off the Antarctic peninsula area: Are killer whales emitting them? *Acoustics Australia*, 45(2), 253–260. <https://doi.org/10.1007/s40857-017-0103-x>
- Rone, B. K., Douglas, A. B., Yack, T. M., Zerbin, A. N., Norris, T. N., Ferguson, E., & Calambokidis, J. (2014). *Report for the Gulf of Alaska Line-Transsect Survey (GOALS) II: Marine mammal occurrence in the Temporary Maritime Activities Area (TMAA) final report*. Retrieved from www.afsc.noaa.gov
- Shaffer, J. W., Moretti, D., Jarvis, S., Tyack, P., & Johnson, M. (2013). Effective beam pattern of the Blainville's beaked whale (*Mesoplodon densirostris*) and implications for passive acoustic monitoring. *The Journal of the Acoustical Society of America*, 133(3), 1770–1784. <https://doi.org/10.1121/1.4776177>
- Shamir, L., Yerby, C., Simpson, R., von Benda-Beckmann, A. M., Tyack, P., Samarra, F., Miller, P., & Wallin, J. (2014). Classification of large acoustic datasets using machine learning and crowdsourcing: Application to whale calls. *The Journal of the Acoustical Society of America*, 135(2), 953–962. <https://doi.org/10.1121/1.4861348>
- Shiu, Y., Palmer, K. J., Roch, M. A., Fleishman, E., Liu, X., Nosal, E. M., Helble, T., Cholewiak, D., Gillespie, D., & Klinck, H. (2020). Deep neural networks for automated detection of marine mammal species. *Scientific Reports*, 10(1), 1–12. <https://doi.org/10.1038/s41598-020-57549-y>
- Sirvic, A. (2015). Variability in the performance of the spectrogram correlation detector for north-East Pacific blue whale calls. *Bioacoustics*, 25(2), 145–160. <https://doi.org/10.1080/09524622.2015.1124248>
- Storrie, L., Lydersen, C., Andersen, M., Wynn, R. B., & Kovacs, K. M. (2018). Determining the species assemblage and habitat use of cetaceans in the Svalbard archipelago, based on observations from 2002 to 2014. *Polar Research*, 37(1), 1463065. <https://doi.org/10.1080/17518369.2018.1463065>
- Thode, A. (2004). Tracking sperm whale (*Physeter macrocephalus*) dive profiles using a towed passive acoustic array. *The Journal of the Acoustical Society of America*, 116(1), 245–253. <https://doi.org/10.1121/1.1758972>
- Thomas, L., & Marques, T. A. (2012). Passive acoustic monitoring for estimating animal density. *Acoustics Today*, 8(3), 35–44.
- Todd, N. R. E., Cronin, M., Luck, C., Bennison, A., Jessopp, M., & Kavanagh, A. S. (2020). Using passive acoustic monitoring to investigate the occurrence of cetaceans in a protected marine area in Northwest Ireland. *Estuarine, Coastal and Shelf Science*, 232, 106509. <https://doi.org/10.1016/j.ecss.2019.106509>
- Trickey, J., Reyes Reyes, M., Baumann-Pickering, S., Melcon, M., Hildebrand, A., & Iniguez, M. (2014). *Acoustic encounters of killer and beaked whales during the 2014 SORP cruise*. Retrieved from <https://www.researchgate.net/publication/284898585>
- Villadsgaard, A., Wahlberg, M., & Tougaard, J. (2007). Echolocation signals of wild harbour porpoises, *Phocoena phocoena*. *Journal of Experimental Biology*, 210(1), 56–64. <https://doi.org/10.1242/jeb.02618>
- Webber, T., Gillespie, D., Lewis, T., Gordon, J., Ruchirabha, T., & Thompson, K. (2021a). *Acoustic detections in Northeast and South Atlantic*. Data downloaded from OBIS-SEAMAP. Retrieved from <https://seamap.env.duke.edu/dataset/2155>
- Webber, T., Gillespie, D., Lewis, T., Gordon, J., Ruchirabha, T., & Thompson, K. (2021b). *Acoustic detections in Indian Ocean*. Data downloaded from OBIS-SEAMAP. Retrieved from <https://seamap.env.duke.edu/dataset/2156>
- Webber, T., Gillespie, D., Lewis, T., Gordon, J., Ruchirabha, T., & Thompson, K. (2022). PAMGuard settings file. *Zenodo*, <https://doi.org/10.5281/zenodo.6558962>
- Wellard, R., Pitman, R. L., Durban, J., & Erbe, C. (2020). Cold call: The acoustic repertoire of Ross Sea killer whales (*Orcinus orca*, type C) in McMurdo Sound, Antarctica. *Royal Society Open Science*, 7(2), 191228. <https://doi.org/10.1098/rsos.191228>
- Whitehead, H., & Weilgart, L. (1991). Patterns of visually observable behaviour and vocalizations in groups of female sperm whales. *Behaviour*, 118(3), 275–296.
- Yack, T. M., Barlow, J., Calambokidis, J., Southall, B., & Coates, S. (2013). Passive acoustic monitoring using a towed hydrophone array results in identification of a previously unknown beaked whale habitat. *The Journal of the Acoustical Society of America*, 134(3), 2589–2595. <https://doi.org/10.1121/1.4816585>
- Yack, T. M., Barlow, J., Roch, M. A., Klinck, H., Martin, S., Mellinger, D. K., & Gillespie, D. (2010). Comparison of beaked whale detection algorithms. *Applied Acoustics*, 71(11), 1043–1049. <https://doi.org/10.1016/j.apacoust.2010.04.010>
- Zilli, D., Parson, O., Merrett, G. V., & Rogers, A. (2014). A hidden Markov model-based acoustic cicada detector for crowdsourced smartphone biodiversity monitoring. *Journal of Artificial Intelligence Research*, 51, 805–827. <https://doi.org/10.1613/jair.4434>

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Webber, T., Gillespie, D., Lewis, T., Gordon, J., Ruchirabha, T., & Thompson, K. F. (2022). Streamlining analysis methods for large acoustic surveys using automatic detectors with operator validation. *Methods in Ecology and Evolution*, 00, 1–13. <https://doi.org/10.1111/2041-210X.13907>