Porpoise Click Classifier (PorCC): A high-accuracy classifier to study harbour porpoises (*Phocoena phocoena*) in the wild

Mel Cosentino[1], Francesco Guarato[1], Jakob Tougaard[2], David Nairn[3], Joseph C. Jackson[1], James F. C. Windmill[1]

[1]*Bioacoustics Group, Centre for Ultrasonic Engineering, Department of Electronic and Electrical Engineering, University of Strathclyde, 99 George Street, G1 1RD Glasgow, United Kingdom*

[2]*Department of Bioscience, Aarhus University, Frederiksborgvej 399, 4000 Roskilde, Denmark*

[3]*Clyde Porpoise CIC,1-1 Allanton Park Terrace, KA29 0AWFairlie, United Kingdom*

Corresponding author: melania.cosentino@strath.ac.uk

Running headline: High Accuracy Porpoise Classifier

Abstract

Harbour porpoises are well-suited for passive acoustic monitoring (PAM) as they produce highly stereotyped narrow-band high-frequency (NBHF) echolocation clicks. PAM systems must be coupled with a classification algorithm to identify the signals of interest. Here, we present a harbour porpoise click classifier (PorCC) developed in MATLAB, which uses the coefficients of two logistic regression models in a decision-making pathway to assign candidate signals to one of three categories: high-quality clicks (HQ), low-quality clicks (LQ), or high-frequency noise (N). The receiver operating characteristics of PorCC was compared to that of PAMGuard's Porpoise Click Detector/Classifier Module. PorCC outperformed PAMGuard's classifier achieving higher hit rates (correctly classified clicks) and lower false alarm levels (noise classified as HQ or LQ clicks). Additionally, the detectability index (*d*') for HQ clicks for PAMGuard was 2.2 (overall *d*' = 2.0) versus 4.1 for PorCC (overall *d*' = 3.4). PorCC classification algorithm is a rapid and highly accurate method to classify NBHF clicks, which could be applied for real time monitoring, as well as to study harbour porpoises, and potentially other NBHF species, throughout their distribution range from data collected using towed hydrophones or static recorders. Moreover, PorCC is suitable for studies of acoustic communication of porpoises.

Keywords: classification, echolocation, logistic regression, NBHF species

I. INTRODUCTION

Studying harbour porpoises *Phocoena phocoena* in their natural environment is a difficult task. They are small and surface for only a few seconds at a time, travelling in groups of three or fewer animals, or as solitary individuals (Hammond et al. 2002). However, they are highly vocal (Linnenschmidt et al. 2013) and are therefore well suited for passive acoustic monitoring (PAM). Harbour porpoises produce highly stereotyped narrow-band high-frequency (NBHF) clicks for echolocation and communication. These clicks have peak and centroid frequencies between 100 and 160 kHz, centred around 130 kHz (Møhl and Andersen 1973), with no spectral energy below 100 kHz (Hansen, Wahlberg, and Madsen 2008). The duration of individual clicks ranges from 50 µs to 175 µs and the half-power (-3 dB) bandwidth is around 15 kHz (Kyhn et al. 2010). Clicks are emitted in series, often referred to as “trains”. A click train is loosely defined as "any series of clicks separated by gradually or cyclically changing inter-click interval suggesting a unit during an echolocation event or a communication signal" (Koschinski, Diederichs, and Amundin 2008). Other odontocetes that produce NBHF clicks are all the porpoises (Phocoenidae), some dolphins of the Lissodelphininae subfamilily, pygmy and dwarf sperm whales (Kogiidae), and the river dolphin *Pontoporia blainvillei* (Galatius et al. 2019).

Different PAM devices are used to study harbour porpoises, including animal-borne devices (Akamatsu et al. 2007), towed hydrophone arrays (e.g., Gillespie et al. 2005; Sveegaard et al. 2011), and static devices (e.g., Carlström 2005; Carstensen, Henriksen, and Teilmann 2006). Static PAM devices can be roughly divided between those that record continuously (e.g., SoundTrap - Ocean Instruments, New Zealand) and click detectors or data loggers that only store information about the transient sounds detected, such as date and time, peak frequency, and amplitude (e.g., C-PODs; Chelonia Ltd., Cornwall, UK). C-PODs, and the earlier version T-POD, are used for a wide variety of studies, including seasonal and geographical changes in distribution (Verfuß et al. 2007) and response to anthropogenic noise (Carstensen, Henriksen, & Teilmann, 2006; Pirotta, Brookes, Graham, & Thompson, 2014). Moreover, they have been used to study porpoise acoustic behaviour (Koschinski, Diederichs, and Amundin 2008), including diurnal variations in echolocation rates and click train patterns (Carlström 2005). Continuous recordings at high sampling rates, required to record harbour porpoise clicks, generate an enormous amount of data and so the storage capacity limits the length of the data collection period. Moreover, data analysis is time consuming and so often part of the data remains unanalysed. On the other hand, click data loggers do not require high storage capacity and thus are more suitable for long-term studies and, since they are coupled with an automatic real-time classifier, the time invested in post-deployment data analysis is reduced significantly. However, it is not possible to carry out post-hoc verification as the C-POD does not record the sound itself, and the detection and classification algorithms are not publicly available.

On the other hand, there are vast amounts of acoustic continuous recordings made using PAM systems that could be used to fill gaps in our understanding of harbour porpoise behaviour and communication in the wild. To that end, however, a classification system that can accurately and reliably identify harbour porpoise clicks is required. A classification system, in simple terms, assigns a given signal x to one of k pre-defined classes according to a series of parameters or functions. For continuous recordings, one of the most used harbour porpoise detector/classifier systems is PAMGuard’s Click Detector and Classifier modules. PAMGuard is a modular, open source software designed to detect and classify marine mammal sounds (Gillespie et al. 2009), and it is used worldwide for a wide range of studies (Cucknell et al. 2016; Lawrence et al. 2016; JNCC 2010). The standard settings of the classifier include a pre-filter (4th order digital Butterworth IIR 10 kHz high pass filter) and a trigger filter (4th order digital Chebyshev IIR 100-150 kHz band pass filter, pass band ripple

2.0). Clicks are classified as produced by porpoises by comparing the test band (110-150 kHz) to control bands (40-90 kHz and 160-190 kHz), with a 6 dB threshold ("general configuration file – porpoise click detection", available at www.pamguard.com). As an open source the software is regularly improved, and although the user can manage the settings, there is no available information about its performance. The precision (i.e., percentage of individual clicks correctly classified as porpoise clicks) reported for an earlier version of this classifier was between 37% and 74%, depending on the settings and background noise, while the proportion of missed clicks was not reported (Gillespie and Chappell 2002). However, the performance of the current version remains unquantified. Additionally, the classifier requires manual verification after the identified clicks have been highlighted and clicks have to be manually selected in echolocation events in order to be extracted for further analysis (Cucknell et al. 2016; Lawrence et al. 2016). Alternatively, many researchers use custom-built classifiers (such as the KERNO classifier of the C-POD), of which neither the algorithm, nor the performance is available. As acoustic recordings continue to accumulate, assessing the performance of available classifiers for comparison purposes and automating these processes becomes essential.

The objective of this study was to develop a harbour porpoise click classifier (PorCC) that improves the performance of existing classifiers, reducing the occurrence of both false alarms and missed clicks, and that provides the user with a simple assessment of the quality of the classified click. PorCC was developed using the output of PAMGuard's Click Detector Module, and uses the coefficients of two logistic regression models to estimate a probability that a given signal was produced by a harbour porpoise. The predictor variables used to build the two logistic regression models were selected because these are the variables most commonly used to describe the temporal and spectral characteristics of harbour porpoise vocalisations (e.g., Kyhn et al., 2013). Once the probabilities are estimated, each signal is assigned to one of three categories: high-quality clicks (HQ), low-quality clicks (LQ), and high-frequency noise (N). These categories were defined based on the characteristics of the waveform, power spectrum, and spectrogram (Fig. 1). The performance of PorCC was tested against manually labelled samples from 18 hours of data collected in two different seasons and in different background noise conditions. Additionally, the performance was tested against PAMGuard's Classifier in a subset of the dataset, consisting of 8 hours of data from one summer day.

II. MATERIAL AND METHODS

A. Data collection

Acoustic data were collected during systematic surveys conducted in the West coast of Scotland, in the Firth of Clyde (55.5254° N, 4.9333° W) during 25 survey days throughout all seasons, between 2016 (n = 20) and 2017 (n = 5), totalling over 210 hours of recordings. Surveys were carried out under sail or engine from the 'Saorsa', a 40-foot sailing vessel. Transect lines were determined in advance and surveyed at a speed between 5 and 7 knots, in different weather conditions, during both day and night times. Surveys were terminated if the sea state reached ≥ 5. No concurrent visual observations were made. Recordings were made using a towed omnidirectional hydrophone array connected to the software PAMGuard (Gillespie et al. 2009) version 1.15.10, and digitised through a St Andrews Instrumentation Ltd. data acquisition card with 16-bit resolution, at a sampling frequency of 500 kHz. The array included two Magrec HP03 hydrophone units, each comprising a spherical ceramic and a HP02 preamp, with a preamp high pass filter set at 2kHz. The hydrophones had a sensitivity of -201 dB re 1V/µPa at 150kHz, and a flat response between 2kHz and 150kHz. The array was towed using a Kevlar-strengthened 100m long cable and the units were 25 cm apart.

PAMGuard's Click Detector Plug-In detects impulsive sounds (i.e., sounds of short duration with abrupt onset and rapid decay) over a given SNR threshold selected by the user (e.g., 6 dB). The detected sound is then saved as an individual audio clip, which also includes a very short recording period before and after the impulsive sound detected. All impulsive sounds detected in a given hour of recording are individually saved in one *.pgdf* file (for PAMGuard Data File) (Gillespie and Oswald 2017). For each audio clip, additional information is attached, such as date and time, time of arrival difference (i.e., delay) with respect to the reference hydrophone, and direction of arrival, estimated using trigonometric methods based on time of arrival differences (Gillespie and Chappell 2002). By extracting individual clips from these files, two datasets were created, one to train PorCC and one to test its performance against manually labelled clips. Additionally, a subset of the testing data was used to compare the performance against PAMGuard's Classifier.

B. Training data

Three categories of signals were defined for the development of PorCC: high quality porpoise clicks (HQ), low-quality porpoise clicks (LQ), and high-frequency noise (N) (Fig. 1). HQ are polycyclic signals with peak frequency between 100 and 160 kHz, no spectral energy below 100 kHz, and duration around 100 µs, matching the description of on-axis harbour porpoise clicks (Au, Kastelein, Benoit-Bird, Cranford, & McKenna, 2006; Hansen et al., 2008). LQ consist of signals slightly different to HQ, for example presenting notches in the power spectrum, or no clear beginning or end of the signal (low signal-to-noise ratio). Noise clips (N) are signals with peak and centroid frequencies between 100 and 160 kHz that do not share other characteristics with harbour porpoise clicks (e.g., oligocyclic, do have energy below 100 kHz).

Of the over 2,500,000 audio clips detected and saved by PAMGuard's Click Detector during the survey period, a subsample of 125,416 (representing 5% of the total) was extracted using a random number generator to ensure they were independent from each other. In order to find good signals to develop the logistic regression models for the classifier (PorCC), an early version was used to assign to each clip a probability of being a harbour porpoise click. Those with high probability (≥ 0.9) were considered to be potential HQ, those with a probability between 0.5 and 0.9 were considered to be potential LQ, and those with a probability < 0.5 were considered to be potential N clips. Subsequently, from these, 5,500 were randomly selected from their respective category to build two logistic regression models, thus 500 were potential HQ, 500 were potential LQ, and 4,500 were potential N. In order to ensure each clip was a good representative of its respective category, all 5,500 clips were then manually verified. Unrepresentative clips were discarded and replaced with clips randomly selected from the original subsample for that particular category.

1. Logistic regression models

In simple terms, the logistic regression model estimates coefficients for each predictor variable in the model and the error term, from which a probability is derived. The predictor variables used were the duration of the signal (estimated as the 80% energy of the clip, µs), peak frequency (PF, kHz) and centroid frequency (CF, kHz), -3dB (BW_{-3dB}, kHz) and root mean square bandwidths (Madsen and Wahlberg 2007), and Q_{RMS} (ratio between CF and BW_{RMS}). Additionally, the ratio between peak and centroid frequencies (Ratio) and the peak value of a cross correlation (XC) performed against a typical harbour porpoise click were used. The click used for the cross correlation was extracted from the original dataset[i], and was selected based on the waveform, power spectrum and spectrogram characteristics and peak-to-peak amplitude (162dB re: 1µPa). Additionally, the waveform was consistent in both hydrophones, and the time of arrival difference between them was 0 (i.e., the orientation of the animal was perpendicular to the array). All predictor variables were explored for normality[ii]. Multicollinearity, that is, when the predictor variables are correlated with each other, was tested using the Pearson χ^2 coefficient and none of the variable pairs had a correlation coefficient higher than ± 0.36, except Q_{RMS} and XC that had a correlation of 0.49[iii].

The response variable for Model 1 is binomial with the outcomes HQ / N and was built using 500 and 4,500 clips of each, respectively. The response variable for Model 2 is also binomial with the outcomes LQ/ N and was built using 500 and 4,500 clips of each, respectively. The same N clips were used for both models. For each logistic regression model, a total of 63 models were tested as a series of reduced models using all possible predictor variable combinations, and the best of each model was identified as the one with the lowest Akaike Information Criteria (AIC) value (Table 1)[iv].

C. PorCC - Classification algorithm

The algorithm of the harbour porpoise click classifier (PorCC) was written in MATLAB 2017a (The Math Works TM, Inc., Natick, MA, USA) and runs on clips previously saved by PAMGuard's Click Detector Plug-In, analysing only those recorded by the first hydrophone on which they impinged. For each clip, the predictor variables identified in the model selection procedure are estimated and two probabilities are calculated using the coefficients obtained from the logistic regression models. Subsequently, a series of if/then statements is applied to assign the clip to one of the three categories previously defined (Fig. 2).

D. Testing data

To test the performance of PorCC, a dataset was created with all clips (n = 265,918) extracted from 5% of *.pgdf* files (i.e., 11 hours of recordings, from ten survey days), which were selected randomly, and all clips (n = 284,231) from the 28th of August 2017 (i.e., eight hours of recordings). Clips with peak and centroid frequencies between 100 kHz and 160 kHz and Q_{RMS}> 4 represented potential harbour porpoise clicks, and so these (n = 70,689) clips were extracted and manually labelled according to the three categories previously defined (Fig. 1), based on the characteristics of the waveform, power spectrum, and spectrogram. The overlap between the training and the testing data was of 442 clips. Subsequently, PorCC was used to classify the clips automatically by estimating the predictor variables and the probability-threshold values of 0.9999 and 0.55 (Fig. 2).

Confusion matrices and receiver operational characteristics (ROC) curves were used to assess the performance of PorCC. The hit rate was calculated for all categories as well as the rate of misclassification (i.e., false alarm and missed clicks) and the precision level. The hit rate is the number of HQ and LQ clips classified as HQ (strict criterion), or as either HQ or LQ (relaxed criterion), divided by the total number of non-N clips. Three analyses were performed: one with only HQ clips, one with only LQ clips, and one with both HQ and LQ clips, all against the N clips. The false-alarm rate is the number of N clips classified as HQ (strict criterion), or HQ or LQ (relaxed criterion) divided by the total number of N clips. In total, this results in six different points of operation in the ROC plot (strict or relaxed criterion combined with HQ, LQ, or both). The missed-clicks rate is defined as 1 minus the hit rate. The precision is defined as the number of clips correctly classified divided by the total number of clips classified into that category.

E. PorCC vs PAMGuard

A subset of the testing dataset for PorCC was used to assess the performance of the porpoise click classifier built-in in PAMGuard and compared it to that of PorCC. This dataset subset contained all clips from the 28th of August 2017 (n = 284,231) of which 30,897 clips had already been manually labelled, having peak and centroid frequencies within the 100-160 kHz range, and Q > 4. PAMGuard's classifier highlights potential harbour porpoise clicks and echoes that the user can manually verify and group into "acoustic events" to later extract them for further analysis. For the purpose of this study, all highlighted clicks were selected without manual verification, assigned to a unique acoustic event, and exported to an SQL database. Putative echoes were included because it was previously noted that this classifier sometimes misidentifies real harbour porpoise clicks as echoes (and vice versa) as well as for comparison purposes, as PorCC also identifies potential echoes (LQ). PAMGuard creates a table within the SQLite database, where information for each of the extracted potential harbour porpoise click is provided, including date, time, and an identification number within the *.pgdf* file where the waveform is saved. Using a custom-built script, and using the identification number, all clips identified by PAMGuard's classifier as potential harbour porpoise clicks were extracted from the *.pgdf* files and saved in a MATLAB structure array for further analysis. Subsequently, clips that were highlighted by PAMGuard's classifier but were discarded by PorCC were manually labelled.

Confusion matrices and receiver operational characteristics (ROC) curves were used to assess the performance of PAMGuard and compare it against PorCC. False alarm, hit rates, and precision levels were also estimated, as well as the detectability index (*d*') (see e.g. Egan, 1975; Tougaard, 2002).

III. RESULTS

A. Logistic regression models

According to AIC values, the best Model 1 (for HQ signals) was that with only Q_{RMS} and duration as explanatory variables, while the best Model 2 (for LQ signals) had five explanatory variables, Q_{RMS}, duration, ratio between peak and centroid frequency, cross-correlation coefficient, centroid frequency, and -3dB bandwidth[v].

B. PorCC performance

PorCC classification process, including estimating all necessary parameters, takes approximately 1 ms per clip. Harbour porpoises produce between fewer than 10 and few hundred clicks per second depending on their behaviour (Clausen et al. 2010; Sørensen et al. 2018; Wright et al. 2017), PorCC shows, therefore, potential for real time application. For HQ, precision was 88.5% (4,475 out of 5,054, 519 of which were LQ and 60 N), false alarm (i.e., N classified as HQ) was 0.0001% (60 out of 537,591 N clips were classified as HQ), and 31.8% of clicks were missed (1,710 were classified as LQ and 382 as N) (Table 2). As precision increases hit rate decreases, that is fewer clicks, of the total available to the classifier, are going to be identified, demonstrating the well-known trade-off between errors: false alarms vs. misses in signal detection and Type I vs. Type II errors in conventional statistics (Fig. 3).

C. PorCC vs PAMGUARD

A total of 30,897 clips from the 28^{th} of August met the criteria for potential harbour porpoise clicks, that is, having peak and centroid frequency between 100 and 160 kHz and Q_{RMS} higher than 4. The results of the comparison of the performance of PorCC (using $Th1 \geq 0.9999$ and $Th2 \geq 0.55$) and PAMGuad's classifier for HQ are shown in Table 2. Based on the detectability indexes (Fig. 3), PorCC outperforms PAMGuard's classifier in all cases, but especially for HQ clicks. The overall precision for HQ for PorCC was 30.8% for

PAMGuard's classifier, assuming that PAMGuard's classifier correctly classified HQ and LQ as such in 100% of the cases, as once clicks are extracted from PAMGuard, there is no information of whether a clip was originally classified as a harbour porpoise click or an echo, which can be considered as equivalent to the HQ and LQ categories[vi].

IV. DISCUSSION

The perfect classifier cannot exist, as detection always will be limited by noise, either external from the environment, or internal. For electronic systems this internal noise is in amplifiers and hydrophones, and for biological systems, this noise will be in the form of spontaneous activity in the neurons. In real-world applications, noise also comes in the form of substantial variation in the temporal and spectral characteristics of acoustic signals. These are affected by many factors, including background noise and the direction from where the signals impinges on the hydrophone, as well as by how the data were collected (e.g., hydrophone own noise, frequency characteristics of the hydrophones) (Richardson et al. 1995). Moreover, in this study, the performance of the classifier is intrinsically linked to the performance of the Click Detector Plug-In in PAMGuard, which in turn depends on the settings selected by the user (e.g., number of samples before and after the signal, SNR thresholds). Despite this, the results of this study show that a classification system based on logistic regression models to identify NBHF vocalisations produced by harbour porpoises outperforms existing classifiers. PorCC can achieve hit rates of over 90% while keeping the false alarm rate below 1% and maintaining high precision levels. The performance of PorCC is expected to be similar, or higher, in data collected using static devices, or in areas with low background noise. Moreover, it has potential for real time application, as it can analyse the equivalent of one hour of data in under 1 minute.

For both logistic regression models, one model was better than the others. It is worth noting that in both cases, the model with the cross-correlation coefficient (XC) as the only explanatory variable appears in the second position after Q, when looking at models with only one explanatory variable[vii]. To classify HQ clicks, cross-correlation analysis, which can be a time costly process, is not necessary and introduces a lot of variation as porpoise clicks are not blueprints of each other. In fact, the cross-correlation coefficient value ranged from 0.0038 to 4.5655, and thus using a threshold in a decision-making process would inevitably include HQ as well as N. The first model containing XC for click detection is fifth on the list. For LQ, on the other hand, XC explains more of the variance in the model, being necessary in the best model, and therefore helps in the classification process. For real-world porpoises, there is likely to be both intra-and inter-animal variation in signals, as well as substantial effects on the frequency spectrum caused by the directionality of the beam and the frequency dependent absorption in the water. This variation is illustrated by the differences in the recorded signals shown in figure 1, most evident by the lack of overlap in frequency spectra of the HQ and the LQ signals. For signals where the parameters are very variable, but where means may be more stable, other types of detectors can be predicted to outperform a cross-correlation receiver. One such receiver is a simple energy detector, which integrates energy within a specified frequency band and a specified duration (Green and Swets, 1966), and this is essentially what the HQ-classifier of PorCC is.

Available data suggest that the variation pattern of inter-click intervals within a click train is indicative of specific behaviours (Clausen et al. 2010; Koschinski, Diederichs, and Amundin 2008; Wisniewska et al. 2018; Sørensen et al. 2018). This is especially true for foraging and feeding behaviour, characterised by inter-click intervals below 10 ms after a phase with much larger inter-click intervals (e.g., Koschinski et al., 2008). PorCC's classification algorithm can be implemented in the output of any transient-sound detector for continuous recordings and, given the low misidentification levels, it is suitable to study the behaviour of wild harbour porpoises, as the variations in inter-click intervals can only be detected if the majority of clicks within a click train are identified. Moreover, these studies can be carried out in data that has already been collected using continuous recordings at an adequate sampling rate, both using towed hydrophone arrays or static devices, such as SoundTrap. The PorCC classification algorithm, including the functions to estimate the different variables and the resulting coefficients, is publicly available and can be coded in other programming languages, such as Python. It could also be incorporated into PAMGuard.

PorCC, like other classifiers, is not exempt of errors, and trying to increase the hit rate would in turn lead to an increase in the false alarm rate, as seen in the change in performance going from a strict to a relaxed criterion. However, the ultimate goal in classification is not to avoid errors, but to manage them. Thus, PorCC provides the user with a general assessment of its performance through the ROC curves, as these show the changes in hit rate with false alarm variations (Tougaard 2002), which results from using different threshold values to classify harbour porpoise clicks. Therefore, users can, *a priori*, manage the level of error according to their needs. Furthermore, depending on the objectives, the user can extract either or both HQ and LQ clicks as well as decide when LQ clicks should be ignored (e.g., single LQ clicks) or taken into account (e.g., studies of click train patterns).

The performance of PorCC for HQ clicks is very high, yet much lower for LQ clicks. This could be the result of some high-frequency noise clips having similar characteristics to LQ clicks, which means the coefficients derived from the second logistic regression model are inefficient to distinguish between LQ clicks and high frequency noise. However, this low performance can also be the result of a level of subjectivity when assigning signals to these categories. This happens to be a fundamental limitation for almost all studies of this kind, where performance of detectors is evaluated on real world data. One must have some means of determining the "true state of the world", i.e. separating signals into those truly originating from porpoises and those that are just random noise. In this study, as in most others, we relied on the superior ability of the human brain to perform pattern recognition in noise and thus measure the performance of the detectors essentially against the performance of a skilled human observer. There is no objective way of determining whether a signal in the array recordings really originated from a porpoise or not. Only under extremely well controlled circumstances, such as when one has a single animal isolated in a pool and a recorder attached to the animal to monitor each and every vocalisation from the animal is it possible to evaluate the absolute detection performance of the detection system and even in such cases, one would suffer difficulties in transferring the experimental settings (limited depth and distance to receiver, training or habituation of the animal etc.) to the situation in real world monitoring.

V. CONCLUSIONS AND FUTURE WORK

The performance of PorCC greatly exceeds that of the currently available classifier in PAMGuard and has potential for real time application as well as to study the acoustic behaviour of harbour porpoises and other NBHF species in the wild, in data collected using both towed hydrophone arrays or static recorders[1]. Future work includes testing PorCC in data obtained using a different recording device (e.g., SoundTrap – Ocean Instruments, New Zealand) and under different survey conditions, and in recordings of harbour porpoises from another population. Additionally, the performance of PorCC will be tested against the performance of C-PODs in data collected simultaneously by a C-POD and a SoundTrap (Sarnocinska et al. 2016).

[1] Preliminary results suggest that PorCC algorithm can be successfully applied to harbour porpoise data recorded with different devices and in different areas, as well as other NBHF species, such as Heaviside's dolphins (*Cephalorhynchus heavisidii*).

ACKNOWLEDGEMENTS

M.C. would like to thank M. Coscarella for his comments and advice during data analysis. Authors would like to thank a P. Ramirez who helped labelling signals in early stages of development of PorCC. Data was collected by community volunteers from Clyde Porpoise CIC with assistance from NA Ventures Trust, Firstport, Bridgend Group and The Hunter Foundation. The research received funding from the Engineering and Physical Sciences Research Council (UK) through grant EP/N509760/1 and via the European Research Council under the European Union's Seventh Framework Programme (FP/2007-2013)/ERC Grant Agreement n. [615030]. J.T. was funded by the Danish Environmental Protection Agency.

AUTHOR CONTRIBUTION

M.C. conceived the ideas, and analysed the data; M.C., J.T., J.F.C.W., and J.C.J. designed the methodology and interpreted the results; D.N. collected the data; M.C. led the writing of the manuscript. All authors contributed critically to the drafts and gave final approval for publication.

COMPETING INTERESTS

The authors have no competing interests declared.

DATA AVAILABILITY

Raw data were generated by the Clyde Porpoise CIC. Derived data supporting the findings of this study and the classification algorithm are available at the Pure Data Repository of the University of Strathclyde.

ORCID

Mel Cosentino https://orcid.org/0000-0001-7837-8669

REFERENCES

Akamatsu, Tomonari, Jonas Teilmann, Lee A. Miller, Jakob Tougaard, Rune Dietz, Ding Wang, Kexiong Wang, Ursula Siebert, and Yasuhiko Naito. 2007. "Comparison of Echolocation Behaviour between Coastal and Riverine Porpoises." *Deep Sea Research Part II: Topical Studies in Oceanography* 54: 290–97. https://doi.org/10.1109/UT.2007.370755.

Au, Whitlow W. L., Ronald A. Kastelein, Kelly J. Benoit-Bird, Ted W. Cranford, and Megan F. McKenna. 2006. "Acoustic Radiation from the Head of Echolocating Harbor Porpoises (*Phocoena Phocoena*)." *The Journal of Experimental Biology* 209: 2726–33. https://doi.org/10.1242/jeb.02306.

Carlström, Julia A. K. 2005. "Diel Variation in Echolocation Behavior of Wild Harbor Porpoises." *Marine Mammal Science* 21 (1): 1–12. https://doi.org/10.1111/j.1748-7692.2005.tb01204.x.

Carstensen, Jacob, Oluf Damsgaard Henriksen, and Jonas Teilmann. 2006. "Impacts of Offshore Wind Farm Construction on Harbour Porpoises: Acoustic Monitoring of Echolocation Activity Using Porpoise Detectors (T-PODs)." *Marine Ecology Progress Series* 321: 295–308. https://doi.org/10.3354/meps321295.

Clausen, Karin Tubbert, Magnus Wahlberg, Kristian Beedholm, Stacy Lynn Deruiter, and Peter Teglberg Madsen. 2010. "Click Communication in Harbour Porpoises *Phocoena Phocoena*." *Bioacoustics* 20: 1–28. https://doi.org/10.1080/09524622.2011.9753630.

Cucknell, A.-C., O. Boisseau, Russell Leaper, R. Mclanaghan, and A. Moscrop. 2016. "Harbour Porpoise (*Phocoena Phocoena*) Presence, Abundance and Distribution over the Dogger Bank, North Sea, in Winter.” *Journal of the Marine Biological Association of the United Kingdom* 97 (7): 1455–65. https://doi.org/10.1017/S0025315416000783.

Egan, J. P. 1975. *Signal Detection Theory and ROC Analysis*. New York: Academic Press.

Galatius, Anders, Morten Tange Olsen, Mette Elstrup Steeman, Rachel A Racicot, Catherine D Bradshaw, Line A Kyhn, and L E E A Miller. 2019. “Raising Your Voice : Evolution of Narrow-Band High- Frequency Signals in Toothed Whales (Odontoceti).” *Biological Journal of the Linnean Society* 126: 231–224.

Gillespie, Douglas, P Berggren, S Brown, I Kuklik, C Lacey, T Lewis, J Matthews, R McLanaghan, A Moscrop, and Nick J. C. Tregenza. 2005. “Relative Abundance of Harbour Porpoises (*Phocoena Phocoena*) from Acoustic and Visual Surveys of the Baltic Sea and Adjacent Waters during 2001 and 2002.” *Journal of Cetacean Research and Management* 7 (1): 51–57.

Gillespie, Douglas, and Oliver Chappell. 2002. “An Automatic System for Detecting and Classifying the Vocalisations of Harbour Porpoises.” *Bioacoustics* 13: 37–61. https://doi.org/10.1080/09524622.2002.9753485.

Gillespie, Douglas, Jonathan C. D. Gordon, Ron Mchugh, David Mclaren, David K. Mellinger, Paul Redmond, Aaron Thode, Phil Trinder, and Xiao Yan Deng. 2009. “PAMGUARD: Semiautomated, Open Source Software for Real-Time Acoustic Detection and Localisation of Cetaceans.” *Proceedings of the Institute of Acoustics*. https://doi.org/10.1121/1.4808713.

Gillespie, Douglas, and M Oswald. 2017. “Binary File Structure for PAMGUARD Detector Output . Version 4 . 0 1 Introduction 2 File Format.”

Hammond, Philip S., Per Berggren, Harald Benke, David L. Borchers, Anne Collet, Mads Peter Heide-Jørgensen, Sonja Heimlich, A.R. Hiby, Mardik F. Leopold, and Nils Øien. 2002. “Abundance of Harbour Porpoise and Other Cetaceans in the North Sea and

Adjacent Waters." *Journal of Applied Ecology* 39: 361–76. https://doi.org/10.1046/j.1365-2664.2002.00713.x.

Hansen, Michael, Magnus Wahlberg, and Peter Teglberg Madsen. 2008. "Low-Frequency Components in Harbor Porpoise (*Phocoena Phocoena*) Clicks: Communication Signal, by-Products, or Artifacts?" *Journal of the Acoustical Society of America* 124 (6): 4059. https://doi.org/10.1121/1.2945154.

JNCC. 2010. "Maintaining Favourable Conservation Status of Harbour Porpoises in UK Water." *Nature*.

Koschinski, Sven, Ansgar Diederichs, and Mats Amundin. 2008. "Click Train Patterns of Free-Ranging Harbour Porpoises Acquired Using T-PODs May Be Useful as Indicators of Their Behaviour." *Journal of Cetacean Research and Management* 10 (2): 147–55.

Kyhn, Line Anker, Frants Havmand Jensen, Kristian Beedholm, Jakob Tougaard, M Hansen, and Peter Teglberg Madsen. 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): 1940–49. https://doi.org/10.1242/jeb.042440.

Kyhn, Line Anker, Jakob Tougaard, Kristian Beedholm, Frants H. Jensen, Erin Ashe, Rob Williams, and Peter Teglberg Madsen. 2013. "Clicking in a Killer Whale Habitat: Narrow-Band, High-Frequency Biosonar Clicks of Harbour Porpoise (*Phocoena Phocoena*) and Dall's Porpoise (*Phocoenoides Dalli*)." *PLoS ONE* 8 (5). https://doi.org/10.1371/journal.pone.0063763.

Lawrence, Joshua M., Eric Armstrong, Jonathan C. D. Gordon, Susan Maersk Lusseau, and Paul Fernandes. 2016. "Passive and Active, Predator and Prey: Using Acoustics to Study Interactions between Cetaceans and Forage Fish." *ICES Journal of Marine Science* 73 (8): 2075–84. https://doi.org/10.1093/icesjms/fst048.

Linnenschmidt, Meike, Jonas Teilmann, Tomonari Akamatsu, Rune Dietz, and Lee A. Miller. 2013. “Biosonar, Dive, and Foraging Activity of Satellite Tracked Harbor Porpoises (*Phocoena Phocoena*).” *Marine Mammal Science* 29 (2): E77–97. https://doi.org/10.1111/j.1748-7692.2012.00592.x.

Madsen, Peter Teglberg, and Magnus Wahlberg. 2007. “Recording and Quantification of Ultrasonic Echolocation Clicks from Free-Ranging Toothed Whales.” *Deep-Sea Research Part I: Oceanographic Research Papers* 54: 1421–44. https://doi.org/10.1016/j.dsr.2007.04.020.

Møhl, Bertel, and S. Andersen. 1973. “Echolocation: High-Frequency Component in the Click of the Harbour Porpoise (*Phocoena Phocoena* L.).” *Journal of the Acoustical Society of America* 54 (5): 1368–79.

Nuuttila, Hanna K., Len Thomas, Jan G. Hiddink, Rhiannon Meier, John R. Turner, James D. Bennell, Nick J. C. Tregenza, and Peter G. H. Evans. 2013. “Acoustic Detection Probability of Bottlenose Dolphins, *Tursiops Truncatus*, with Static Acoustic Dataloggers in Cardigan Bay, Wales.” *The Journal of the Acoustical Society of America* 134 (3): 2596–2609. https://doi.org/10.1121/1.4816586.

Pirotta, Enrico, Kate L. Brookes, Isla M. Graham, and Paul M. Thompson. 2014. “Variation in Harbour Porpoise Activity in Response to Seismic Survey Noise.” *Biology Letters* 10: 20131090. https://doi.org/10.1098/rsbl.2013.1090.

Richardson, John W., R. Greene, Jr. Charles, Charles I. Malme, and Denis H. Thomson. 1995. *Marine Mammals and Noise*. https://doi.org/10.1016/S0022-0981(96)02706-2.

Sarnocinska, Joanna, Jakob Tougaard, Mark P. Johnson, Peter Teglberg Madsen, and Magnus Wahlberg. 2016. “Comparing the Performance of C-PODs and SoundTrap/PAMGUARD in Detecting the Acoustic Activity of Harbor Porpoises (Phocoena Phocoena).” In *Proceedings of Meetings on Acoustics. Fourth International*

Conference on the Effects of Noise on Aquatic Life, 27:070013. https://doi.org/10.1121/2.0000288.

Sørensen, Pernille Meyer, Danuta Maria Wisniewska, Frants Havmand Jensen, Mark P. Johnson, Jonas Teilmann, and Peter Teglberg Madsen. 2018. “Click Communication in Wild Harbour Porpoises (Phocoena Phocoena).” *Scientific Reports* 8: 9702. https://doi.org/10.1177/0964663912467814.

Sveegaard, Signe, Jonas Teilmann, Per Berggren, Kim Nørgaard Mouritsen, Douglas Gillespie, and Jakob Tougaard. 2011. “Acoustic Surveys Confirm the High-Density Areas of Harbour Porpoises Found by Satellite Tracking.” *ICES Journal of Marine Science* 68 (5): 929–36. https://doi.org/10.1093/icesjms/fsr025.

Tougaard, Jakob. 2002. “Signal Detection Theory, Detectability and Stochastic Resonance Effects.” *Biological Cybernetics* 87: 79–90. https://doi.org/10.1007/s00422-002-0327-0.

Verfuß, Ursula K., Christopher G. Honnef, Anja Meding, Michael Dähne, Roger Mundry, and Harald Benke. 2007. “Geographical and Seasonal Variation of Harbour Porpoise (*Phocoena Phocoena*) Presence in the German Baltic Sea Revealed by Passive Acoustic Monitoring.” *Journal of the Marine Biological Association of the UK* 87: 165–76. https://doi.org/10.1017/S0025315407054938.

Wisniewska, Danuta Maria, Mark P. Johnson, Jonas Teilmann, Ursula Siebert, Anders Galatius, Rune Dietz, and Peter Teglberg Madsen. 2018. “High Rates of Vessel Noise Exposure on Wild Harbour Porpoises (*Phocoena Phocoena*) Can Disrupt Foraging.” *Proceeding of the Royal Society B* 285: 20172314. https://doi.org/10.1098/rspb.2017.2314.

Wright, Andrew J., Tomonari Akamatsu, Jonas Teilmann, Kim Nørgaard Mouritsen, Signe Sveegaard, and Rune Dietz. 2017. “Silent Porpoise: Potential Sleeping Behaviour Identified in Wild Harbour Porpoise.” *Animal Behaviour* 133: 211–22.

https://doi.org/10.1016/j.anbehav.2017.09.015.

Table 1. Series of logistic regression models for Model 1 and Model 2. Only the best five are shown here. See text for description of the variables used[viii]. The outcomes of the response variable for Model 1 are high-quality harbour porpoise clicks or high frequency noise, and for Model 2 are low-quality harbour (LQ) porpoise click or high-frequency noise (N). AIC = Akaike's Information Criterion.

ID	**Predictor Variables – Model 1**	**ΔAIC**
1	Q_{RMS} + Duration	0
2	Q_{RMS} + Duration + Ratio	1.64
3	Q_{RMS} + Duration + BW	1.67
4	Q_{RMS} + Duration + CF	1.78
5	Q_{RMS} + Duration + XC	1.96

ID	**Predictor Variables – Model 2**	**ΔAIC**
1	Q_{RMS} + Duration + Ratio + XC + CF + BW	0
2	Q_{RMS} + Duration + Ratio + XC + BW	1.19
3	Q_{RMS} + Duration + Ratio + CF + BW	19.19
4	Q_{RMS} + Duration + XC + CF + BW	20.07
5	Q_{RMS} + Duration + Ratio + BW	20.87

Table 2. Confusion Matrices. Comparison of correct and misidentification levels between PorCC and Porpoise Click Detector/Classifier Module in PAMGuard, and overall performance of PorCC. HQ = high-quality harbour porpoise clicks. LQ = low-quality harbour porpoise clicks. Noise = high and low--frequency noise (i.e., anything that is not a porpoise click).

Labelled		**PorCC**		**PAMGuard**			**PorCC**	
	Total	**HQ**	**Noise**	**HQ**	**N**	**Total**	**HQ**	**Noise**
HQ	1833	564*	1269[¥]	1209	113[¥]	6567	4475	382
Noise	965	477+1601[#*]	279,355	25[#]	280,034	537,591	60	533,228

*Of the total of 3,017 clips highlighted by PAMGuard as potential harbour porpoise clicks, 1,601 had Q_{RMS} < 4 and peak and centroid frequencies outside of the 100 and 160 kHz range, therefore they were not captured by PorCC, as they were discarded at the first step.

[#] False alarm (N clips classified as HQ clicks divided by the total number of N clips)

[¥] Missed clicks (HQ clicks classified as N divided by the total number of HQ click).

Figure Captions

Figure 1. Examples of the categories defined to develop the harbour porpoise click classifier (PorCC). a) High-quality harbour porpoise click (HQ). b) low-quality harbour porpoise click (LQ). c) high-frequency noise (N). Wigner plot (centre plot), waveform (lower plot), and power spectrum (right plot).

Figure 2. Flowchart illustrating the decision-making pathway of the harbour porpoise click classifier (PorCC). CF = centroid frequency. PF = peak frequency. Th = probability thresholds. Prob = Probability.

Figure 3: Receiver operating characteristics (ROC) curves. Dots represent false alarm rates and hit rates associated with detection of HQ-clicks (solid black line), LQ-clicks (black dashed line) and both types combined (grey line), all against a background of N-clicks. Top figures show performance of PAMGuard. Curves are best fitting ROC-curves, generated under the assumption of Gaussian underlying distributions with equal variance. Bottom figures show performance by PorCC under two different criteria: strict (only clicks classified by PorCC as HQ) and relaxed (all clicks classified as either LQ or HQ). Figures to the left and right contain same data, but right figures are plotted on double probit (probability) axes.

---

[i] See supplementary material at [URL will be inserted by AIP] for a figure (Fig. S1.1) of the click used as a model against which a cross-correlation is performed in the PorCC classification algorithm.

[ii] See supplementary material at [URL will be inserted by AIP] for the histogram and distributions of all variables used to develop the logistic regression models (Fig. S2.1).

[iii] See supplementary material at [URL will be inserted by AIP] for a correlation plot (Fig. S2.2) of all variable pairs.

[iv] See supplementary material at [URL will be inserted by AIP] for a complete list of all logistic regression models performed to develop the PorCC classification algorithm (Tables S1 and S2).

[v] See supplementary material at [URL will be inserted by AIP] for a complete list of all logistic regression models performed to develop the PorCC classification algorithm (Table S1 and S2)

[vi] See supplementary material at [URL will be inserted by AIP] for examples of signals misclassified by both PorCC (Fig. S3.1 to S3.6) and by PAMGuard's Click Classifier Module (Fig. 3.7).

[vii] See supplementary material at [URL will be inserted by AIP] for a complete list of all logistic regression models performed to develop the PorCC classification algorithm (Tables S1 and S2)

[viii] See supplementary material at [URL will be inserted by AIP] for a complete list of all logistic regression models performed to develop the PorCC classification algorithm for high-quality signals (Table S1).