STATISTICAL CLASSIFICATION OF ODONTOCETE CLICKS

Douglas Gillespie and Marjolaine Caillat

Sea Mammal Research Unit, Gatty Marine Laboratory,
University of St. Andrews, Fife, KY16 8LB, Scotland

ABSTRACT

To the best of our knowledge, all odontocetes produce some kind of click like vocalisation, which is used primarily for echolocation but may also play a role in social communication. Characteristics of these echolocation pulses range from the broad band but relatively low frequency clicks of sperm whales to the ultrasonic, narrow-band clicks of harbour porpoise. Although these clicks are often easily detected, it can be difficult to classify them to species, thereby hampering efforts to monitor and study odontocetes using passive acoustics. Candidate clicks from three species were detected using a simple energy trigger, operating in the frequency band of interest. The clicks were then identified to species using two different statistical classifiers to separate beaked whale vocalisations from those of other odontocete sounds. In the first, a number of parameters (peak frequency, mean frequency, sweep frequency, click duration, width of principal spectral peak and the relative energy in different frequency bands) were calculated and a tree classifier was used to separate clicks of different species. In the second, the spectral energy in 32 relatively coarse energy bands 1.5 kHz wide were used as input to a multivariate classifier. Both classifiers were trained and tested using data provided to the 3rd International Workshop on Detection and Classification of Marine Mammals using Passive Acoustics in order to assess the classifiers performance with Blainville's beaked whales, short-finned pilot whales and Risso's dolphin clicks. The methods were also applied to survey data collected using a towed hydrophone deployed from a sailing research vessel in the Bahamas. Some of the towed hydrophone data were collected over the US Navy's AUTEC range where independent confirmation of beaked whale vocal activity was available from bottom-mounted hydrophones.

SOMMAIRE

L'état actuel de nos connaissances nous permet d'affirmer que tous les odontocetes émettent des sons de type impulsifs, aussi appelés clics, destinés surtout à l'écholocation, mais ils peuvent également être utilisés pour la communication. En fonction des espèces, ces clics peuvent couvrir une bande de fréquence plus ou moins large. Le cachalot produit des clics couvrant une large bande de basses fréquences, alors que chez le marsouin, l'écholocation est caractérisée par des clics ultrasoniques couvrant une bande de fréquence étroite. En sélectionnant les clics ayant une puissance supérieure à un certain seuil avec un simple détecteur d'énergie dans la bande de fréquence qui nous intéressait, nous avons collecté les clics de 3 espèces (Baleine à bec de Blainville, globicéphale tropical et dauphin de Risso). Deux méthodes d'analyse nous ont permis de discriminer les sons de la baleine à bec de Blainville de ceux des 2 autres espèces. Pour la première méthode, différent paramètres (pic de fréquence, fréquence moyenne, variation de fréquence, durée du signal, largeur du spectrogramme et énergie relative dans les différentes bandes de fréquences) ont été extraits de chaque clic et utilisés dans un arbre de classification afin de séparer les espèces. Pour la seconde méthode, l'énergie contenue dans 32 bandes de 1.5kHz a servi de données pour une analyse multivariée. Les 2 classificateurs ont été entrainés et testés en utilisant les données de la 3ième commission internationale de détection et de classification des mammifères marins en utilisant l'acoustique passive. L'objectif était de mesurer la performance des classificateurs pour discriminer la baleine à bec de Blainvilles, le globicéphale tropical et le dauphin de Risso. Ces 2 méthodes ont ensuite été appliquées sur des données collectées au Bahamas à partir d'hydrophones tirés par un voilier de recherche. Quelques données furent collectées au dessus de la zone du réseau d'hydrophone Autec, appartenant à la marine Américaine, permettant d'obtenir une confirmation indépendante de l'activité sonore des baleines à bec via ce réseau sous-marin.

1. INTRODUCTION

Concern over the link between the use of military sonar and strandings of beaked whales has led to much research into the acoustic behaviour of beaked whales in recent years. Archival recording tags (Johnson and Tyack, 2003) attached to Blainvilles Beaked Whales (*Mesoplodon densirostris*) and Cuviers Beaked Whales (*Ziphius cavirostris*) (Johnson *et al*, 2004, Zimmer *et al*, 2005, Madsen *et al*, 2005) show that they produce narrow-band clicks with pulse lengths of around 200 μs and most of the energy concentrated in the 25 to 40 kHz energy band.

Practical applications for the management of risks to beaked whales require that methods be developed to detect them more efficiently than can currently be achieved using visual observers (Barlow and Gisiner, 2006). Passive acoustic monitoring can potentially be used to assess the distribution and abundance of beaked whales and has also been proposed as a method for detecting animals in the immediate vicinity of vessels using sonar. Real time mitigation requires that beaked whale clicks be detected with a high efficiency, although it may not be necessary to identify to species level. Surveys of abundance do not require that detection efficiency be high, only that it be known, although species identification is more important than it might be for mitigation.

In this paper we demonstrate how beaked whale clicks may be detected and how they may be statistically separated from clicks from short-finned pilot whales (*Globicephala macrorhynchus*) and Risso's dolphins (*Grampus griseus*). Pilot whales produce both tonal vocalisations (whistles) and clicks (Weilgart and Whitehead, 1990). Both the whistles and clicks can be heard by humans since they are at lower frequency than beaked whale clicks. Risso's dolphins on the other hand echolocate at much higher frequencies. Risso's clicks are generally broad band, having energy between 30 and 100 kHz and an average 3 dB bandwidth of 39.7 kHz (Philips *et al*, 2003). The Risso's data analysed in this study were only sampled at 96 kHz and had been low-pass filtered at 38 kHz, and were therefore only acquiring the lower frequency components of the Risso's dolphin clicks.

2. METHODS

Detection and classification algorithms were trained and tested on data provided to the 3rd International Workshop on Detection and Classification of Marine Mammals using Passive Acoustics, which contained clicks from Blainvilles's beaked whales (BBW), short-finned pilot whales (SFP) and Risso's dolphins (RD). All of these data were in the form of wav file recordings, sampled at 96 kHz and containing data from a single bottom-mounted hydrophone either at the Bahamas Atlantic Undersea Test and Evaluation Center (AUTEC) range (BBW and SFP) or from the Southern California Operating Range (SCOR) in San Clemente Island (RD), California, USA.

Data were also available from a towed hydrophone array deployed from a sailing vessel undertaking line transect surveys in the Bahamas. Some of these data were collected at the AUTEC range, where animals were being simultaneously monitored on bottom-mounted hydrophones.

Most of the analysis was done using RainbowClick software (www.ifaw.org/sotw). RainbowClick was originally developed for the detection and analysis of sperm whale echolocation clicks (Gillespie, 1997). As well as containing click detection and classification algorithms, and applying them either to real time data or archived data from file, RainbowClick provides the user with an interactive display where detected clicks may be easily selected and their waveforms and spectra examined by the user. Clicks, or groups of clicks, can be exported via a database for more detailed analysis by other software packages (e.g. Matlab).

2.1. Click Detection

Detection and classification of clicks were conducted as clearly separate stages of the analysis. In the click detection stage, regions of sound files found to have significant energy in the 25 – 40 kHz band were extracted and stored.

Clicks were detected using an algorithm operating on time series data. The algorithm is designed for real time operation, using infinite impulse response filters (IIRF) for efficient data analysis (Lynn and Fuerst, 1989).

Optimal detection of beaked whale clicks is achieved by band-pass filtering the data in the 25 to 40 kHz range. However, the statistical classifiers (see below) require a comparison between acoustic energy within the 25 to 40 kHz beaked whale band and acoustic energy at lower frequencies. The detector therefore contains two separate filters as shown in Figure 1. The use of two filters allows the detector to operate only on signal within the band of interest, but data used in the classification stage can use signal in a wider band. Removal of low-frequency data, particularly at frequencies with wavelengths on the order of or longer than the clip length, is essential to avoid large sidelobes dominating the spectra used in the classifiers. The first filter was a second-order high-pass Butterworth with a cut-off frequency of 7 kHz and the second had a fourth order band-pass 25 kHz to 40 kHz Chebyshev response.

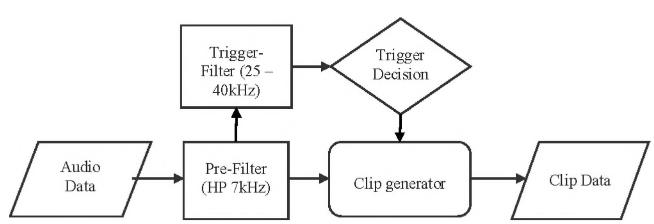

Figure 1. Schematic diagram of click detection process.

Training set extraction

Candidate clicks were detected in all samples of the training data set. To avoid occasional clicks from non-target species, or noise entering the training set, an operator (Caillat) then examined click files using RainbowClick in order to select a sub-sample of clicks which appeared to have consistency of amplitude and inter click interval with other clicks in the data (i.e. appeared to form part of an echolocation click train) and had waveform and spectral properties consistent with published literature for the three species. Numbers of clicks selected to form the training sample were BBW 6399, SFP 1555 and RD 609.

2.2. Click Classification – Method 1

Parameter Extraction

Beaked whale clicks are characterised by having most of their energy in the 25 to 40 kHz band. More detailed spectral analysis using a Wigner-Ville (WV) distribution shows that there is in fact a slight upsweep in frequency during a typical beaked whale click (Johnson *et al*, 2006 and **Figure 2**).

Six parameters were measured for each candidate click from the detector. From the power spectrum, the mean frequency, the peak frequency and the ratio of the acoustic energy in the 25 to 40 kHz band compared with that in the 10 to 25 kHz band were measured.

From the WV distribution, the sweep of the click, the maximum 'bandwidth' of the click at any point along its length and the length of the click were extracted. This was achieved by taking the maximum point in the WV distribution and then performing a regional search around it in order to establish a contour 6dB below the energy at the maximum (**Figure 2**). The 'ridge' of maximum acoustic energy along the length of the click was also extracted. The bandwidth was taken as the maximum distance between the lower and upper edges of the contour and the length of the click as the time between its start and its end. The sweep was taken as the difference from the start to the end of the maximum energy ridge. This broadly follows a pattern of parameter extraction found to be useful in detecting and classifying right whale contact calls (Gillespie, 2004).

Classification

The six parameters described above were computed for all clicks in the training data set. A classifier was then realised using a tree classification function (Breiman *et al*, 1993). Tree classifiers divide data into groups using multiple binary splits. Each split uses a single variable or parameter to divide the data into two groups, the variable and the split value at each node being chosen to maximise the deviance between the two groups. This process is repeated until every group contains perfectly homogeneous data (i.e. clicks of only one type). For practical classifiers, the number of nodes is limited (the tree is 'pruned') to avoid problems of over-fitting to training data.

In order to establish which of the six parameters extracted for each click were most useful in classification, 11 different classification models were tested, each using a different sub selection of the six parameters (**Table 1**). Trees were pruned to five nodes. The training data were split and 2/3 of the

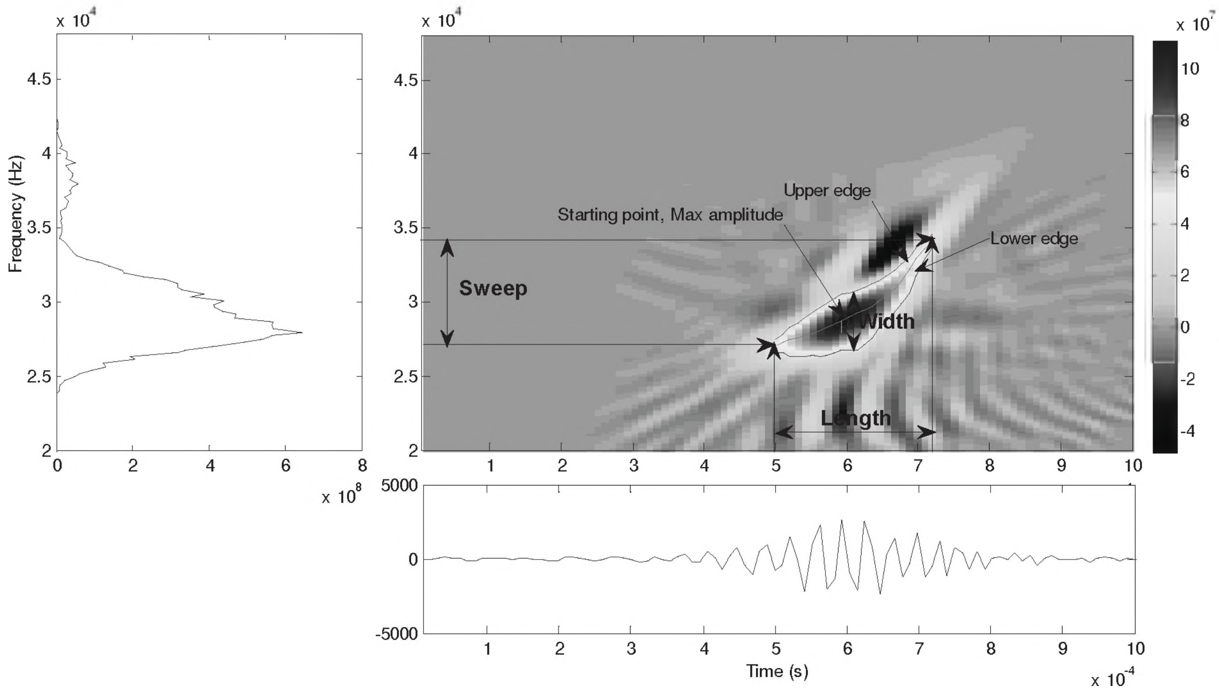

Figure 2 Wigner-Ville distribution, power spectrum and waveform for a typical beaked whale click showing the –6dB contour around the click and the ridge of maximum energy.

Table 1. Parameter selection for eleven different models tested using the tree classifier.

Parameter	Model										
	1	2	3	4	5	6	7	8	9	10	11
Length	√					√		√			√
Width	√	√				√		√			√
Sweep	√	√	√				√	√			√
Mean Frequency	√	√	√	√			√		√	√	√
Peak Frequency	√	√	√	√	√				√		√
Energy Ratio	√	√	√	√	√	√	√	√		√	

data randomly selected for classifier training. Equal numbers of clicks for each species were selected from the remaining 1/3 of the data for testing. Since we had the fewest clicks from Risso's dolphins (609) this meant that 203 clicks from each species were used in the testing samples. Bootstrapping was used to train and test each model 500 times, each time using a different random sample of training and test data and the average error rate from the 500 bootstraps taken.

2.3. Click Classification – Method 2

Parameterisation for the second classification method simply took the power spectrum for each click and divided it into coarse energy bins 1.5 kHz wide. A 512 point FFT was calculated for each click (each click clip either being truncated or padded with zeros to achieve the correct length) and the relative energy in bins 1.5 kHz wide (8 FFT bins) taken, i.e. if $S(\omega)$ is the power spectrum, then the coarse spectrum $S'(\omega)$ is

$$S'(\omega_i) = 10\log_{10}\left(\frac{\sum_{8i}^{8(i+1)} S(\omega_i)}{\sum S(\omega)}\right) \qquad \text{Equation 1.}$$

giving a total of 32 parameters for each click, although only those above 7.5 kHz were used in classification.

Clicks were classified using a one-way Multivariate Analysis of Variance (one way MANOVA) to divide training data into groups (Krzanowski, 1988.). The MANOVA calculates a linear discriminant function chosen to maximise the separation between groups and produces an matrix of eigenvalues which can be used to calculate canonical variables which are simple linear combinations of the parameters (relative energies in 1.5 kHz bands). These canonical variables can then be used to assign clicks to different groups using a relatively small number of variables.

Although such a classifier should work equally well, or better, with finer-scale data, the limited size of the training sample might have made the model over fitted to the available data if all 256 frequency bins had been used as input parameters.

3. RESULTS

3.1. Click Classification – Method 1

Distribution plots for the six parameters extracted for each species are shown in **Figure 3**. Most of the distributions overlap heavily, particularly for beaked whales and Risso's dolphins. Pilot whale clicks are generally at lower frequency than those of the other species, making them stand out on plots of mean and peak frequency as well as the energy ratio.

Results of the tree classification are shown in **Figure 4**. When all three species are analysed together, the error rate varies between approximately 20 and 30 % for the pruned tree. If only beaked whale and pilot whale clicks are included in the analysis, the error rate is extremely low.

In both cases, the worst performing model is model 9, which does not include the energy ratio (**Table 1**) thereby indicating that this is one of the more important parameters. Of particular interest for practical applications are models 4, 5, 9 and 10, which only use parameters extracted from the power spectrum. Computation of the Wigner Ville distribution takes approximately 2000 times longer than the calculation of a power spectrum, so any method that does

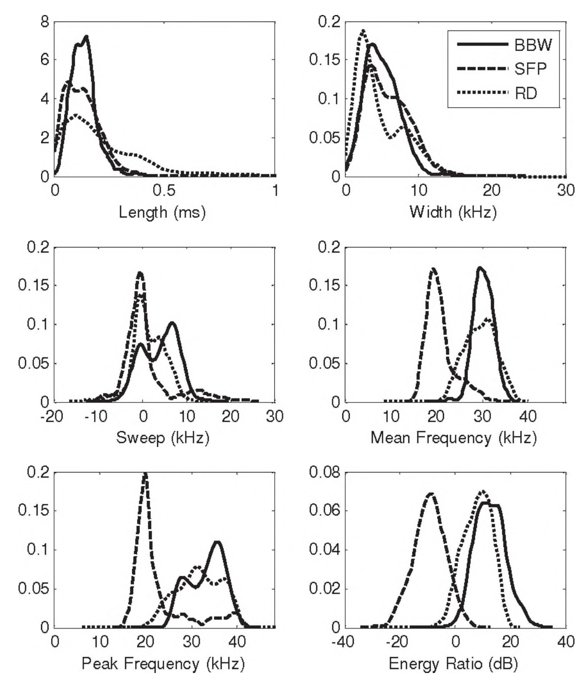

Figure 3: Distributions of the six parameters for each species.

not require those parameters will be much easier to implement in real time systems.

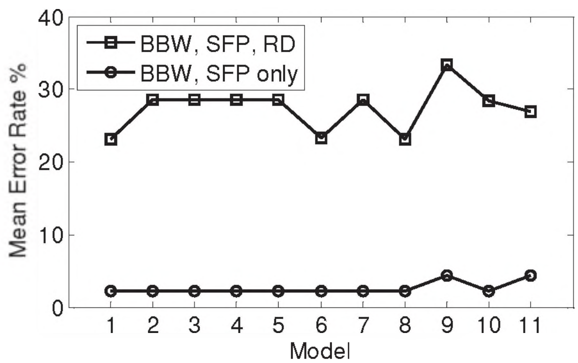

Figure 4: Tree classification error rates for the 11 models.

3.2. Click Classification – Method 2

Figure 5 shows the mean coarse spectra for different click types. **Figure 7** shows distributions of the first two canonical variables from the MANOVA analysis of coarse spectra. Clearly, there is good separation between the three species when this method is employed although there is still slight overlap between beaked whales and pilot whales.

4. APPLICATION TO 'UNKNOWN' WORKSHOP DATA

The 'Unknown' workshop data were analysed using the second classification method since it had better overall performance for all three species than the first. During analysis, it was assumed that only the three species present in the training set were present. If C1 and C2 are the first two canonical variables (**Figure 7**) then clicks were classified according to the following selection criteria:

BBW if C1 > -0.5 and C2 < 2
else
SFP if C1 < -1.5 and C2 < 2
Else
RD if C2 > 3
Else
Click is unclassified.

Numbers of clicks of each type in each of the test files are shown in **Table 2**. Since the classifier was trained only to identify three species, spotted dolphins (*Stenella attenuata*) and sperm whales *(Physeter macrocephalus*) were misclassified as pilot whales. Clicks from both these species have their predominant energy below 20 kHz (Lammers *et al* 2003; Møhl *et al*, 2003) so this misclassification is not surprising.

Click waveforms and spectra were viewed with RainbowClick. Of particular interest are the small numbers of beaked whale clicks in data sets 3 and 4. These beaked

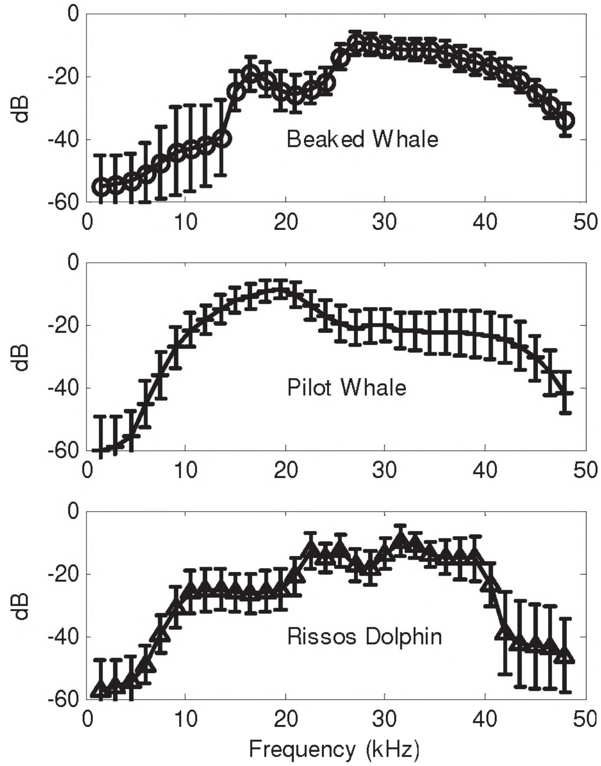

Figure 5: Mean coarse spectra for the second click parameterisation.

whale clicks appeared in short bursts of 3 – 4 clicks and are quite clearly from beaked whales. However, this is not the case for the clicks classified as beaked whales in other files which, on visual inspection of waveforms and spectra, are clearly false classifications.

5. APPLICATION TO BAHAMAS TOWED HYDROPHONE DATA

As well as being tested with the workshop dataset, the detector and classifiers were tested using data collected using a towed hydrophone deployed from the sailing research vessel 'Odyssey' while undertaking line transect surveys around the Bahamas in June and July, 2007. Recordings were made at a sample rate of 192 kHz.

The second classification method did not work at all well with the towed hydrophone data, classifying large numbers of false triggers from vessel noise as beaked whales. The tree classifier on the other hand did perform well and picked out a number of beaked whale click trains, some of which were coincident with clicks being detected on the bottom-mounted hydrophones at AUTEC. In all, 172.5 hours of recordings were analysed. Initial processing to detect and classify clicks took approximately one week. It then took an operator (Gillespie) one day to go through the data and confirm beaked whale detections. A typical click from the towed hydrophone is shown in **Figure 6**.

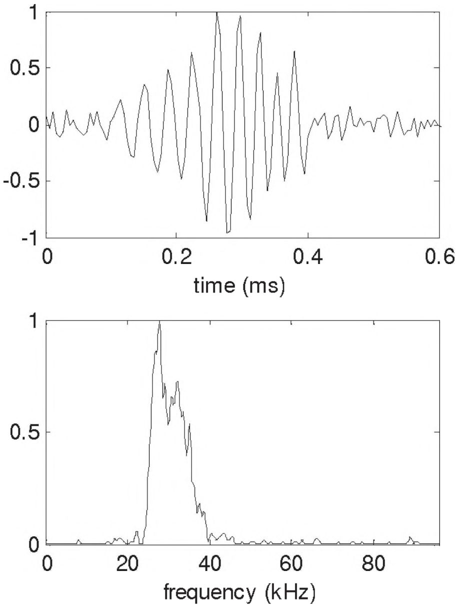

Figure 6: Waveform and power spectrum of a beaked whale click recorded on the towed hydrophone.

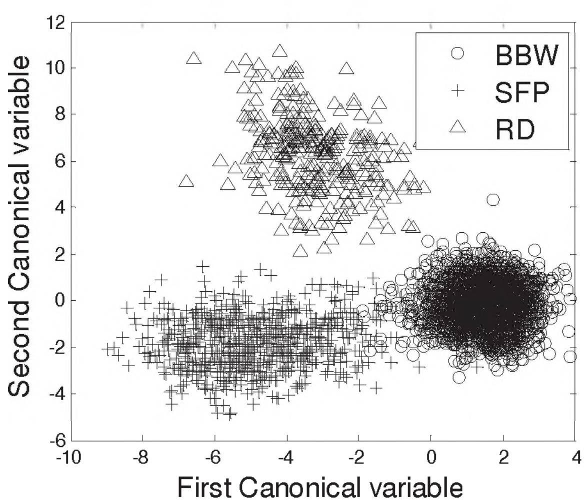

Figure 7: distributions of the first two canonical variables from the MANOVA analysis.

6. DISCUSSION

Clicks from beaked whales, pilot whales and Risso's dolphins can be detected and separated using statistical classifiers operating on parameters extracted from the clicks power spectra. Two methods have been tested, the first of which performed less well with the test data from bottom-mounted hydrophones than the second method, but was more stable when applied to data collected on a towed hydrophone array.

Table 2: Clicks identified within the 'unknown' test data files.

Set	Number of clicks				Truth
	BBW	SFP	RD	No Class	
1	1502	531	5	542	BBW+SFP
2	1566	1000	5	131	BBW
3	11	15200	0	24	Spotted*
4	13	16674	0	28	Spotted*
5	13	462	10879	466	RD
6	947	1217	2	30	BBW
7	154	104	3414	758	RD
8	994	4749	5	2391	SFP
9	0	4909	1	30	Sperm

*Visual inspection of these data following analysis with RainbowClick shows clicks which clearly matched published waveforms and spectra for BBW.

Using the first method, when all three species were analysed together, the error rate was around 20 to 30 % for the pruned tree. Although this appears high, if multiple clicks were analysed together (as they generally are) this might improve. However this depends on whether or not the errors are being caused by random noise (in which case we would expect an improvement) or genuine overlap in click parameters in which case the improvement may be small since adjacent clicks from the same whale are likely to be very similar.

One of the reasons why the first method did not perform particularly well at separating BBW and RD clicks is probably that the RD echolocate at higher frequencies than could be represented in recordings sampled at 96 kHz (Philips et al, 2003), so the workshop data sets only contained the very low edge of the RD spectra. Had higher frequency recordings been available, it is likely that separation of these species would have been as straightforward as the separation of SFP and BBW. Given that many researchers are able to collect data at a sample rate of only 96 kHz, asking whether or not it is possible to separate BBW and RD in such data is a valid question. Unfortunately, however, the RD data had been collected at a different location and heavily filtered above 38 kHz, so if there was a difference, it was lost in these data. Conversely, and perhaps more worrying is the possibility that the detectors somehow 'tuned in' to this fundamental difference in the recordings rather than differences in the sounds themselves.

When applied to the 'unknown' data set, the detector found beaked whale like clicks in a number of recordings. Visual inspection of waveforms, power spectra and inter-click intervals convinces us that some of these are genuine beaked whale clicks. Although no visual sightings of beaked whales were made, these species are notoriously difficult to spot and it is quite possible that some were present, even though they were not seen.

A fundamental problem with classifiers trained to detect only a small number of species is that they will tend to misclassify data from any other source (be it noise from a vessel or some other cetacean species). For classifiers of the type presented here to be genuinely useful, they must be trained with data from all species and noise sources likely to be present in the data to which they are to be applied.

Although beaked whales have been successfully detected on bottom-mounted hydrophones, beaked whales echolocate only when undertaking deep foraging dives (Tyack *et al*, 2006). Since their clicks are produced in a narrow, forward-facing beam (Zimmer *et al*, 2005), it has therefore been suggested that detection using towed hydrophones close to the surface is unlikely. When applied to data collected using a towed hydrophone, the detector and classifier were able to pick out several beaked whale click trains. This result is extremely encouraging and opens up the possibility for towed hydrophone surveys for beaked whales. However, further work is required to establish with what efficiency this can be achieved as a function of detection range.

As well as separating beaked whale clicks from those of dolphin species (SFP and RD), some applications, such as abundance estimation, may require the separation of species within the beaked whale family. Although Johnson *et al*, (2004), show clear differences between clicks of Blainville's and Cuvier's beaked whales, vocal behaviour of many other beaked whale species remains largely undocumented. Another research priority for the coming years is therefore to obtain broadband recordings from other beaked whale species.

The click detector has been implemented into the PAMGUARD open source software (www.pamguard.org) and work is underway to implement the classification methods.

7. ACKNOWLEDGEMENTS

This work was funded by the Joint Grants Scheme operated by the Defence Science and Technology Laboratory and the Natural Environment Research Council. We wish to thank the organisers of the 3rd International Workshop on Detection and Classification of Marine Mammals using Passive Acoustics for the provision of the training data sets and Dianne Claridge, Jonathan Gordon and other participants in the Bahamas Marine Mammal Research Organisation surveys for the towed hydrophone data. RainbowClick software development was funded by the International Fund for Animal Welfare to promote benign and non-invasive research.

8. REFERENCES

Barlow, J., and Gisiner, R. 2006. Mitigating, monitoring, and assessing the effects of anthropogenic sound on beaked whales. J. Cetacean Res. Manage. **7**.3, 239–249.

Breiman, L., Friedman, J., Olshen, R., Stone C. 1993. *Classification and Regression Trees*, Chapman and Hall, Boca Raton.

Gillespie D. 1997. An acoustic survey for sperm whales in the southern ocean sanctuary conducted from RSV *Aurora Australis*. Reports of the International Whaling Commission, 47, 897–906.

Gillespie, D. 2004. Detection and classification of right whale calls using an `edge' detector operating on a smoothed spectrogram. Canadian Acoustics 32(2): 39-47

Johnson, M., and Tyack, P. L. 2003. A digital acoustic recording tag for measuring the response of wild marine mammals to sound. IEEE J. Ocean. Eng. 28, 3–12.

Johnson, M., P. T. Madsen, W. M. X. Zimmer, Aguilar. de Soto, N. and P. L. Tyack. 2004. Beaked whales echolocate on prey. Proceedings of the Royal Society of London Series B-Biological Sciences 271:S383-S386.

Johnson M., Madsen P.T., Aguilar de Soto N., Tyack P. 2006 Foraging Blainville's beaked whales (*Mesoplodon densirostris*) produce distinct click types matched to different phases of echolocation, J. Exp. Biol., 209 5038-5050.

Krzanowski, W. J. 1988, *Principles of Multivariate Analysis*. Oxford University Press.

Lammers, M., Au, W., Herzing, D. 2003. The broadband social acoustic signaling behavior of spinner and spotted dolphins. J. Acoust. Soc. Am. 114(3), 1629-1639.

Lynn, P., and Fuerst, W. 1989. *Introductory Digital Signal Processing with Computer Applications*. John Wiley & Sons,New York

Madsen P.T., Johnson M., Aguilar de Soto N. 2005 Biosonar performance of foraging beaked whales (*Mesoplodon densirostris*) J. exp. Biol. 208, 181-194.

Møhl, B., Wahlberg, M., Madsen, P.T., Heerfordt, A. & Lund, A., 2003. The monopulsed nature of sperm whale clicks. J. Acoust. Soc. Am, 114, 1143–1154.

Philips, J. Nachtigall, P., Au, W., Pawloski, J., Roitblat, H. 2003. Echolocation in the Risso's dolphin, *Grampus griseus*. J. Acoust. Soc. Am. 113 605-616

Tyack P.L., Johnson M., Aguilar de Soto N., Sturlese A., Madsen P.T.M.2006. Extreme diving of beaked whales, J. Exp. Biol., 209. 4238-4253

Weilgart, L.S. and Whitehead, H. 1990. Vocalizations of the North Atlantic pilot whale (*Globicephala melas*) as related to behavioral contexts. Behav. Ecol. Sociobiol. 26: 399-402.

Zimmer, W., Johnson, M., Madsen, P., Tyack, P. 2005. Echolocation clicks of free-ranging Cuvier's beaked whales (*Ziphius cavirostris*). J. Acoust. Soc. Am., 117(6) 3919-3927