Bio-Duck Activity in the Perth Canyon. An Automatic Detection Algorithm.

David Matthews (1), Rod Macleod (1) and Robert D. McCauley (2)

(1) Maritime Operations Division, DSTO, HMAS Stirling, Rockingham, WA Maritime Operations Division, DSTO, HMAS Stirling, Rockingham, WA
(2) Centre of Marine Science and Technology (CMST), Curtin University, GPO Box U 1987, Perth 6845, WA

Abstract

Recently analysed data from Curtin University has revealed a significant amount of "bio-duck" activity in the Perth Canyon during December 2002. The name "bio-duck" originates from sonar operators on board the old Oberon class submarines who thought that the sound resembled that of a duck. Surprisingly this is not the case for the Curtin data. The difference however may be due to onboard audio processing prior to the operators hearing the sound that was absent in the Curtin data. It should also be noted that for both data sets the origin of the sound is unknown.

For the recent data there exists two distinct types of call. One long period (T~3.1 sec) and one short period (T~1.6 sec) covering the frequency range 60 Hz < f < 1000 Hz. This could have major implications on the operations of some of the sonar on-board the Collins class submarines. Consequently an algorithm was written to automatically detect the presence of bio-duck. In order to eliminate the effect of amplitude variations between data files a signal-time ratio method was used for a third octave band centred around 125 Hz. The reliability of this algorithm was estimated by comparing it's output with that of the manual analysis of 2240 data files (23 days). It was found to have a 93% success rate in detecting the bio-duck. This will allow quick analysis of large amounts of data to investigate annual variations and also give a method for automatic detection on board the submarines. The results will be discussed.

Introduction

In order to achieve the best possible tactical advantage it is important for the RAN Submarine Fleet to have a good knowledge of it's littoral underwater environment. While parameters such as the sound speed profile (SSP) and sea bed parameters have obvious effects on sonar performance large effects can also be caused by local biological noise. Owing to the shear size of the Australian littoral coastline the variation in this noise can change drastically from location to location. A detailed knowledge of this variation would give a better prediction of it's variability and the possibility of using it to an advantage. Initial work has begun in the Perth Canyon with Curtin University recording data in the trench over a prolonged period of time. This work needs to be extended to other areas of strategic interest around Australia

One example of biological noise recently recorded in the Perth Canyon is that of the "bio-duck". Despite being well know to the submarine community (especially the old Oberon class submariners) very little is known about it's origin, maximum speed or diving depth. Recent data recorded by Curtin Uni [1] has revealed a substantial amount of bio-duck activity during the months of October to December 2002. Owing to the large quantity of data collected it was necessary to develop some auto-detection algorithm to analyse the data. However, one problem with testing such algorithms is in estimating the effectiveness and false alarm rates. Our solution to this was to manually analyse 23 days of data (2240 files) using audio and spectragram techniques to create a data base by which we could then benchmark our algorithm against.

Experimental data

Curtin University have been collecting a large amount of data in the Perth Canyon since November 2001 as part of their Western Australia Exercise Area Blue Whale Project [1]. In this work they have been deploying data loggers at regular periods throughout the year at various locations in the Perth Canyon. As mentioned, this has resulted in a substantial amount of useful information not only about Blue Whales but about any other marine noise that may be also present at the time of data collection (in the frequency band of the data collection loggers). For the work reported in this paper one of these data sets was used (set 2612 from mooring #9). This data logger was deployed in 450 m of water at the position 115^{o}00.11'E, 31^{o}53.75'S. Data was recorded at a sampling rate of 4 kHz every 15 minutes for a duration of 120 seconds. The data covered the dates from 14/10/02 to 20/12/02 and consisted of 6451 records. The first 23 days (2240 records) of this data was analysed manually to build a data base for algorithm testing.

Bio-Duck Characteristics

Figure 1 shows a typical example of a bio-duck signature. The white graph shows the time domain signal and has been superimposed on the frequency domain spectrogram for comparison. The most obvious characteristic is the extremely reparative nature of the signal. Most of the energy is located in the frequency band 50 to 300 Hz however for more intense signals harmonics have been observed up to 1 kHz similar to those shown in figure 1. This data set also contained a large amount of humpback activity as indicated on the figure.

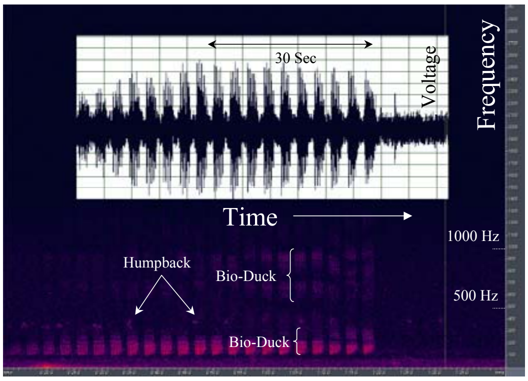

Figure 1.

Typically signals have been found to last for as little as 2 seconds up to the entire file size of 2 minutes.

From the data analysed two main types of bio-duck signature have been observed. These are shown in Figure 2. The first type (Type 1) consists of two frequency sweeps of duration ~ 0.5 seconds occurring every 1.6 seconds. Most of the energy is concentrated in the frequency band 50- 200 Hz. The second signal type (Type 2a) consists of a burst of 5 sweeps lasting for approximately 2.1 seconds followed by a quiet period for 1 second. The frequency span of the type 2 signals is slightly smaller than that of the Type 1 ranging from 100 to 200 Hz.

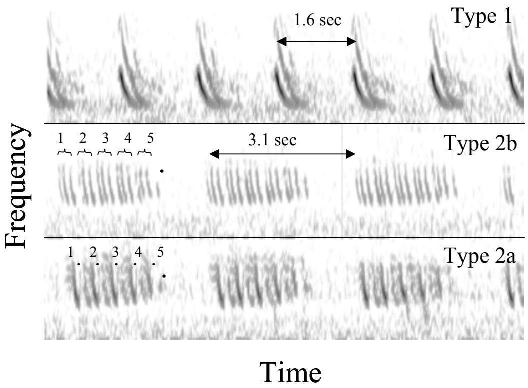

Figure 2.

The Type 2b signatures however have a similar structure as the Type 2a except that each of the primary lines has been split into 3 separate lines. These characteristics have been used to design an auto-detection algorithm described below.

Auto-Detection Algorithm

As mentioned above the signal level of bio-duck calls in various files had large variations. Consequently it was not possible to use a simple threshold method for detecting their presence. To overcome this a signal-time ratio method was used and is shown graphically in Figure 3.

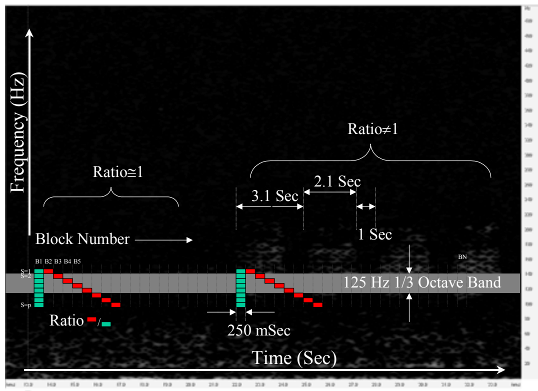

Figure 3

All data files were down sampled to 2 kHz and filtered using a 1/3 octave filter centred at 125 Hz. This centre frequency was chosen as both the type 1 and type 2 bio-duck signals have a significant amount of energy in this band. Comparison of this third octave band with others also adds to the versatility of the algorithm and gives some control in eliminating false alarms from sources such as electrical noise. This has not been implemented in this paper.

The data file in question was then divided into 250m s blocks and the ratio calculated between the average value of a certain block and that of the average value of it's nearest neighbor (Δt=0). This ratio was calculated for the entire file and the result averaged to produce a value for Δt=0. The time difference between blocks was then increased to 250 ms (Δt=250 ms) and the same process repeated to obtain the average block ratio for Δt=250 ms. This was repeated for all values of Δt up to 9 seconds in increments of 250 ms. The result is an array of 36 points that represents the average value of the ratio of the blocks from 0 to 9 seconds. The block size and increment size of 250 ms was chosen specifically for the bio-duck detection however it could be tailored to suit any reparative waveform.

Referring back to figure 3 and applying the procedure described above. For two adjacent blocks with similar energy in both, the ratio will be approximately 1. As the blocks pass through the region containing bio-duck there will be a very short time when the ratio ≠ 1 but overall the average value of all the ratios will be close to 1. As the distance between blocks is increased (Δt increased) there will be larger periods of time when the ratio between blocks ≠ 1 and the average value of the total ratio will be greater than or less than one. When Δt⇒ the repetition rate of the bio-duck signal (either 1.6 second for Type 1 or 3.1 seconds for Type 2) then the total ratio will once again ⇒ 1. Therefore for a file containing a type 2 bio-duck signal the total ratio should go through 3 maxima over the 9 seconds. This is indeed the case as shown in Figure 4.

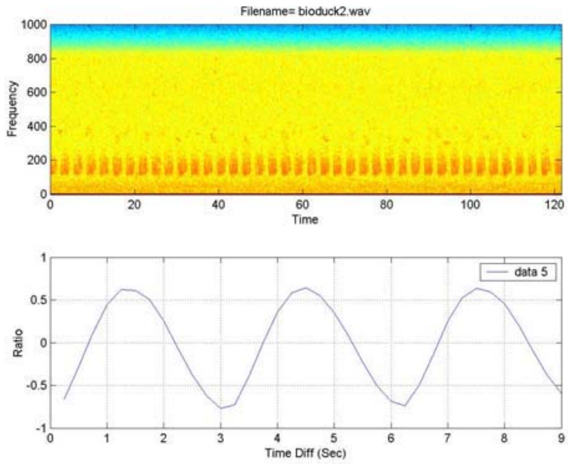

Figure 4

The top graph in Figure 4 shows a spectragram of a loud type 2 bioduck calling for the entire duration of the file (120 seconds). The bottom graph shows the average of the block ratios using the procedure described above. This clearly shows the period of 3 seconds as expected. In order to check the method the same process was used on a file that contained no bio-duck signal but did contain noise from a passing merchant ship (which produces a significant amount of noise in the same frequency band as the bio-duck). This is shown in Figure 5. In this case the ratio shows no periodicity similar to that observed in Figure 4.

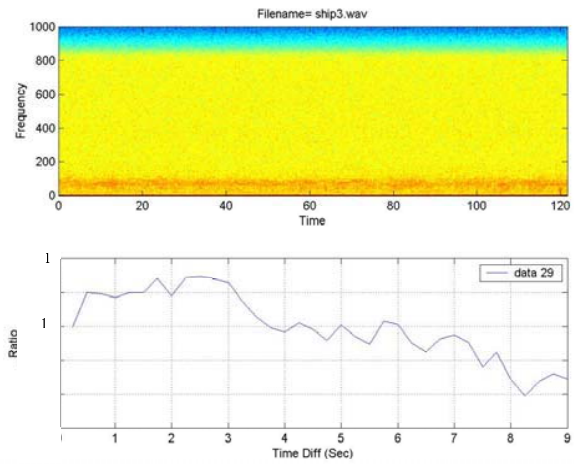

Figure 5

The periodicity T observed in Figure 4 was estimated by fitting a sinusoid to the data using a least squares curve fit routine from the Optimisation Toolbox in MatLab. If $2.8 \le T \le 3.3$ then type 2 bio duck is detected. If $1.5 \le T \le 1.7$ then type 1 bio-duck is detected.

Results

As mentioned above, in order to make an estimate of the effectiveness of the algorithm, the first 23 days of data was analysed manually (2240 files). Unfortunately at the time of this analysis we were not aware of the various type of bio-duck and consequently both types have been classed as one. Figure 6 shows a comparison of the manual results (Blue) to that produced by the algorithm (Red).

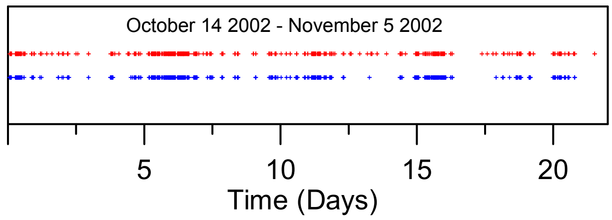

Figure 6

As can be seen there is reasonably good agreement between the two results. From the 2240 files analysed the manual method detected 399 occurrences of bio-duck. Using the same data the algorithm found 371. This is a 93% detection rate. However, closer observation of Figure 4 shows that there are some differences between the two output files indicating that the algorithm produces a false alarm rate ~ 8%. Reexamination of some of these files has indicated that the manual method had missed some bio-duck activity (usually when shipping noise was present as well as bio-duck). Some problems have also been found when the signal to noise ratio is very small. Despite this the occurrence of miss detection appears to be very low.

The remaining files in this set were then analysed using the algorithm and the output is shown in Figure 7. Using the algorithm it was possible to distinguish between both types of bio-duck as indicated on the plot.

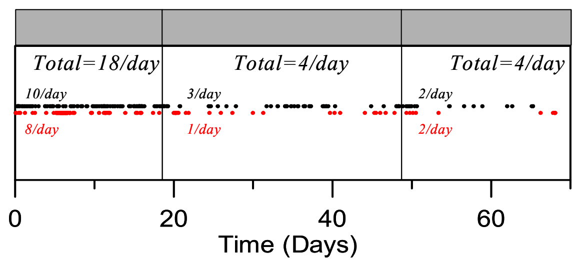

Figure 7

The black circles represent the type 2 and the red circles the type 1 bio-duck. From the data shown in Figure 7 it can be seen that the highest bio-duck activity is observed in October with the numbers dropping off in the next two months. For October the number of bio-duck calls was found to be 18/day which drops off to approximately 4/day for November and December.

Conclusions

An algorithm has been developed to automatically detect the presence of bio-duck noise. The reliability of this algorithm was estimated by comparing it's output with that of the same data analyzed manually. It was found to have a 93% success rate of detecting bio-duck. Estimations of the false alarm rate and missed detections are still under investigation. By using a time ratio method for detecting the bio-duck signal it was possible to distinguish between two different type of call. One with a low repetition rate (T=1.6 seconds) and one with approximately double the period (T=3.1 seconds). The algorithm was then used to analyse the remainder of the data set giving an estimate of the bio-duck activity in the Perth Canyon during the months of October to December 2002. We intend to use it to analyse more data recorded in the same location at different times. This will give us an indication of any annual variation.

Acknowledgments

Thanks goes to the following people: To Matt Legg and Cheryl Smith-Gander for help in the coding and Damien Killeen for useful comments during development of the algorithm. To Christopher Donald, Boyd Wykes, Colin Trinder, Doug Cato and Steve Cole from Australian Defence who have financed and encouraged data collection for this project, for which we are most grateful. The expertise of the crews of various Perth lobster boats, particularly Paul Pettorini and crew, have been essential for deploying and recovering gear.

References

[1] Western Australian Exercise Area Blue Whale Project, Rob McCauley, John Bannister, Chris Burton, Curt Jenner, Susan Rennie and Chris va Etten, Milestone 4, September 2003. CMST Report R2003-33, Project – 350.