FINAL REPORT
6th International workshop on the Detection, Classification, Localization and Density Estimation of Marine Mammals

Dr Douglas Gillespie
Sea Mammal Research Unit,
Scottish Oceans Institute,
University of St Andrews,
St Andrews, KY16 8LB
Scotland
phone: +44 1334 462663 fax: +44 1334 463443 email: dg50@st-andrews.ac.uk

Award Number: N000141310262
http://soi.st-andrews.ac.uk/dclde2013

LONG-TERM GOALS

The efficient detection classification and localization of marine mammals has in recent years, become critical to many Navy activities. Much of this need has been driven by the well-publicized mass stranding's of beaked whales in response to the use of Navy tactical mid frequency sonar (e.g. Cox et al., 2005) but there are many other areas in which Navy activities have the potential to negatively impact on marine mammal species including vessel collisions and live fire exercises. In order to better mitigate against the harmful effects of these activities and to help ensure the long term health of marine mammal populations, the Navy requires efficient methods for the detection of marine mammals in order to:

a) Survey for Marine mammals so that areas of low density can be identified in order that certain activities can be sited in these low density areas.
b) Provide real-time mitigation immediately prior to and during potentially harmful activities so that they can be curtailed if animals are in the vicinity.
c) Survey for marine mammals in order to determine that population impacts have not occurred.

While it is unlikely that Passive Acoustic Monitoring (PAM) will ever fully replace visual monitoring, in recent years it has started to play an increasingly important role, particularly for the detection of deep divers such as beaked whales which are almost impossible to detect visually (Barlow and Gisiner, 2005) and for long term monitoring where maintaining vessels and visual survey teams at sea can become prohibitively expensive.

Our long term goals are therefore to:

a) Improve on methods for passive acoustic Detection, Classification, Localisation and Density Estimation of marine mammals.
b) Understand the limitations, whether fundamental or technical of passive acoustic methods.

OBJECTIVES

This challenge is being addressed through a series of biennial meetings of internationally recognised experts in the field. The meetings firstly encourage an exchange of knowledge between leading experts in the field. Through analysis of a common annotated dataset related to a current topical problem, in order that different developers algorithms and techniques can be directly compared.

This award was used to part fund the 6th of these biennial meetings, which was held in St Andrews, Scotland between 12 and 15 June, 2013. Approximately half of the award was used to fund staff time during the preparatory stages of the meeting, the other half to fund travel and workshop participation for students in order to help encourage young people entering this field of research.

APPROACH

Organisation of the meeting followed a well established system of talks, discussion, and poster sessions in which people could present and discuss their work. Meetings were held over four days, with approximately 4.5 hours of oral presentations each day and an hour of poster viewing. At the end of each session, all speakers were encouraged to participate in group discussion chaired by an expert in that field. These discussion sessions were vital in testing ideas and assisting in guiding future work.

WORK COMPLETED

The workshop passed off successfully in June. 118 people attended in total from 11 different coutries (Australia, Canada, France, Germany, Italy, Japan, The Netherlands, Portugal, Spain, The United Kingdom and The United States). The workshop website remains live at http://soi.st-andrews.ac.uk/dclde2013/ and all abstracts ara available for viewing in the abstact book which can be downloaded from the website.

RESULTS

Three of the four days of the workshop started with an invited speaker. These were Dr Vincent Janick from the University of St Andrews, who spoke about “Call variation in marine mammals”, Tiago Marquez who gave a talk titled “Have you heard about it? The perfect passive acoustic density estimation survey is out there and we are looking for it!” and Dr Sofie van Parijs, from NOAA Fishers who spoke about “Using DCLDE for management and conservation of North Atlantic right whales.” In addition there were 62 other oral presentations from delegates and 33 poster presentations.

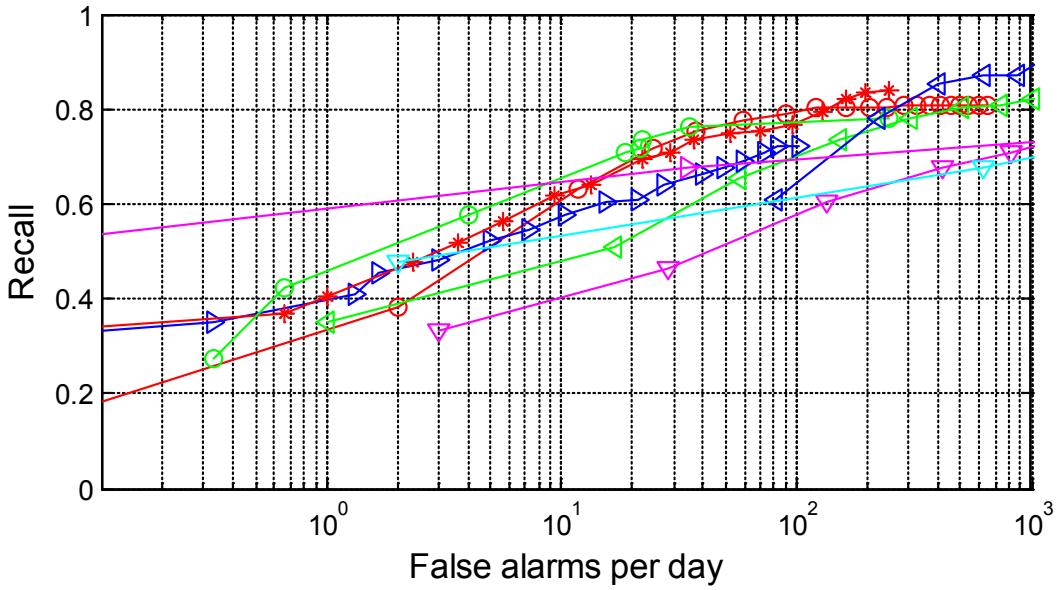
Common dataset:
For the 2013 workshop a dataset has was created which brought us back to where we started in 2001 – the detection of right whales. A great deal has been accomplished since 2001, but right whales remain one of the most endangered great whales on the planet. North Atlantic right whales continue to suffer ship strikes and fishing gear entanglement and is possible that they will become impacted by the installation of renewable energy devices in the coming years. In the Southern hemisphere right whales have been dying in unprecedented numbers at Peninsula Valdes, Argentina.

Whereas at the 2001 workshop we were only able to provide a few dozen right whale sounds we now, thanks to the enormous efforts of researchers at NMFS and Cornell BRP, have thousands of carefully

annotated sound files including gunshots as well as upsweeps. Of course anyone who works on detecting marine mammals knows that detecting the sounds you want is only half the problem. Particularly in the case of right whales, NOT detecting sounds from other species in the same area is a task of equal or greater magnitude. We therefore made sure that the datasets include plenty of data without any right whales at all (so far as we can tell).

The dataset was released in two stages, the first several months prior to the workshop, complete with annotation data so that delegates could train their algorithms. The second part was released only weeks before the meeting so that people could run their algorithm on the data, but not (or hopefully not) have time to retune and over-train their algorithm for that specific dataset.

As with previous meetings only a small number of participants examined the workshop dataset in detail. However, we were able to compare results from nine different algorithms. We do not have permission from all participants to include their results here, but the following is a summary plot of algorithm performance for calls with an SNR of between 6 and 9 dB.



What is striking about these results is that the performance of many of the algorithms is very similar, with a detection efficiency (Recall) of around 50% for a few false alarms per day. Furthermore, one of the results, the red stars, is an algorithm developed using the training set form the first DCLDE workshop in 2003 (Gillespie, 2004). This may indicate that there is a fundamental limitation in the data – that many right whales calls just sound like humpbacks and only a subset (around 50%) can ever reliably be assigned as right whales.

Discussions
During the meeting, there was also considerable discussion in the following topic areas:

Workshop proceedings:
A peer reviewed special issue of an acoustic journal has been produced for each of the previous five DCLDE workshops. However, it was noted that at the time of the workshop the proceedings for the previous 2011 workshop had not yet been published and many people felt that by submitting their papers for inclusion in that special issue, it had significantly delayed publication since the process had only been able to proceed at the speed of the slowest contribution. It was therefore agreed that a special issue would not be generated for the 2013 DCLDE meeting. However, several delegates stated that it was a good thing to have so many DCLDE related papers in a single volume and that special issues can be useful reference documents. John Hildebrand and colleagues offered to look into online publishing options prior to the next DCLDE workshop, which will probably be organised by their Whale Acoustic lab in San Diego in 2015.

Future Training sets:
The development of algorithms is often hampered by the lack available ground truth data which makes it impossible to tell which algorithms are working better and which are not. Just about all training sets have their limitations, in particular:

- Vocalisations from a particular species can change geographically, between sub populations and over time, so an algorithm optimised for one dataset may not work so well with other data
- Background noise types also vary considerably between locations, so an algorithm which performs well on a training set with only one or two types of background noise may perform badly when a different noise is introduced.
- Most training sets only contain data for a single species, whereas most real world situations contain multiple species.
- No training data sets yet exist which can be used to ground-truth density estimation methodologies.

There are a number of ways in which ground-truthed data sets can be produced, all of which have their drawbacks and limitations.

Human Annotation is an effective way of marking out call types for many species of marine mammal and can be useful when the humans ability to pick out sounds far exceeds the computer algorithms. However, as the algorithms performance starts to approach that of a human, then the algorithm will start to pick out sounds that the human missed and these will then be falsely flagged as false alarms.

Synthetic Data. If synthetic sounds are mixed with synthetic noise, then at least we know the truth in the training files. Similarly, real animal sounds can be mixed at known times into files of background noise. However, it is unlikely that either of these types of synthetic data can truly capture the variability of sound types that can occur in real data.

Tag Data. If tag data are available it is often possible to clearly identify and annotate high signal to noise ratio sounds on the tag. If hydrophones are also present, then the tag data can be used as a ground truth for the hydrophone data. A drawback of this is the small number of species and noise scenarios for which tag data are available.

It was agreed that a working group should be formed to take discussion of datasets further and an email will be sent to all workshop participants to see who is interested in participating in such a group.

Creating Interest Through Competitions:
It was suggested that the community may become more motivated and may also draw in new people if a prize were offered for the most effective algorithm to address a specific problem. Cornell BRP has recently been trying this for the detection and classification of right whale sounds through Kaggle (http://www.kaggle.com/c/the-icml-2013-whale-challenge-right-whale-redux). Whether this generates a useful new algorithm or something very over trained for that specific data set, waits to be seen. The organisers of the next DCLDE workshop must decide whether or not to incorporate a competition element into the meeting.

IMPACT/APPLICATIONS

Without a formal proceedings, it is difficult to assess impact in a quantitative way. However, most leading experts in this field agree that these meetings are essential in exchanging ideas and moving the field forwards. We are hopeful that a number of particpipants will be publishing on the topic of DCLDE in the coming months.

REFERENCES

Barlow, J., and Gisiner, R. (2005). "Mitigating, monitoring and assessing the effects of anthropogenic sound on beaked whales," Journal of Cetacean Research and Management 7, 239.

Cox, T. M., Ragen, T. J., Read, A. J., Vos, E., Baird, R. W., Balcomb, K., Barlow, J., et al. (2005). "Understanding the impacts of anthropogenic sound on beaked whales," Journal of Cetacean Research and Management 7, 177.

Gillespie, DM 2004, ' Detection and classification of right whale calls using an 'edge' detector operating on a smoothed spectrogram ' *Canadian Acoustics* , vol 32, no. 2, pp. 39-47.