**1st Working Group Meeting on the Multi-study OCean acoustics Human effects Analysis (MOCHA)**

University of St Andrews, St Andrews
15 and 16 February 2012

## 1. Summary

This document briefly reports on the first meeting of the US Office of Naval Research sponsored MOCHA working group, held on $15^{th}$ – $16^{th}$ February 2012 in St Andrews. The overall aim of the group is to develop and implement innovative statistical methodologies for the analysis of behavioral response study (BRS) data. In this start-up meeting, we discussed ways to integrate data from existing and ongoing BRS projects as well as combine methods across species, and defined initial technical groups to work on specific methodologies.

## 2. Background and objectives of working group

The US Navy has long recognized the need to quantify the impact of active sonar at the level of individual animal responses and also in terms of population consequences. It has therefore invested heavily in studies designed to quantify and mitigate impact.

Behavioral response studies (BRSs, also sometimes called controlled exposure experiments or CEEs) are experiments aimed at directly quantifying the relationship between potential anthropogenic disturbances and their effect on specific marine mammals. The first US Navy-funded BRS focusing on mid-frequency active sonar was conducted in 2007 and 2008 in the Bahamas with the aim of collecting baseline data on animal behavior, and conducting CEEs to measure responses to different sound stimuli. A major focus of the Bahamas work was beaked whales, which are thought to be some of the species most vulnerable to sounds from sonar exercises. The 3S project, which aimed to determine behavioral response of killer, sperm and pilot whales to different sonar frequencies, commenced in 2006 with funding from The Royal Norwegian and Dutch Navies, and in 2008 the additional US-Naval support allowed a significant expansion of the effort. This project was conducted in Norwegian waters and is now being followed by the $3S^2$ project (starting 2011), which plans to expand the range of species studied and investigate the effectiveness of ramp-up as a mitigation strategy. The SOCAL-BRS experiment conducted its first season of data collection in 2010. SOCAL-BRS is a five-year study in southern Californian waters, which aims to expand on the work conducted in the Bahamas by increasing the number of species studied and by integrating with other field efforts being conducted in the same region. Recently, funding has been secured to carry out CEEs on cetaceans off Cape Hatteras and Hawaii, in order to investigate their response to the sounds of potential predators that resemble the sounds of military sonar, specifically testing the influence of species-specific social structure on those responses.

BRSs are an important approach for studying the short-term response of animals to potential stressors and offer a direct measure of the effect of sound on behavior. In a BRS, focal animals are identified and their behavior monitored using visual observations, passive acoustics and attachment of animal-borne tags. A sample of animals is exposed to a stimulus, such as a potentially disturbing sound sequence, and their response monitored. Diverse experimental protocols have been employed, such as escalating the dose until a response is observed, using various stimuli, exposing the same animal multiple times and varying the context of the experiment (such as the animal's behavior before the experiment or the path of the source vessel). Care is taken to ensure the experimental animals are not unduly stressed, for example by ceasing the stimulus as soon as a response is detected, or if any potentially dangerous responses are observed. Various measurements are recorded, including location (in 3D) through time, acoustic behavior, behavior at the surface, etc.

Typically the first stage of analysis of BRS data is to synthesize the observational data to determine whether there was a response or not (or the magnitude of any response) and at what level of dose. Secondly, results from individual exposures may be analyzed together to estimate a (context-specific) dose-response function. Both stages are fraught with practical difficulties. At the first stage, it is not clear how best to combine the diverse input metrics (such as position, orientation, swim speed, dive time, social configuration and surfacing behavior) all measured through time, into a single measure of response, or indeed whether this is desirable. At both stages, analyses need to account for dependencies in the data, because many measurements are taken in close succession on the same subjects. The same subjects may also be exposed multiple times. While there are many measurements on each subject, there are few exposures in total, and often even fewer individual subject animals. Analysis approaches based on traditional statistical hypothesis testing and estimation have typically been employed; however these have restrictive assumptions and their statistical power is often very low. This has the undesirable effect that only studies on species showing the clearest responses are likely to provide statistically significant results that are easily publishable. This can lead to a potential bias in published effect levels, such that animals are thought to be more sensitive on average than is actually the case.

The overall objective of this project is to develop and implement innovative statistical methodologies for the analysis of behavioral response study data. The focus is on studies funded by the US Navy to investigate the effect of mid-frequency active sonar. These include the Sirena sonar trials on sperm whales, Bahamas BRS, SOCAL BRS, Norwegian 3S and 3S2, Cape Hatteras EK60 experiments and Cape Hatteras and Hawaiian killer whale BRS. A working group format was adopted, bringing together researchers undertaking BRS studies and statisticians specializing in the analysis of biological experiments of this kind with the aim to maximize the inferences that can be drawn from current and ongoing studies as well as to provide advice on future studies. This approach enables us to complement and enhance the analytical work already being undertaken, as well as to be flexible and incorporate new ideas as they arise in working group sessions.

The project has four specific objectives.

1. Improve methods for combining diverse behavioral measures into metrics of behavioral response. Consideration will be given to obtaining metrics that can be linked to biological consequences. The group will consider the strengths and limitations of the various possible behavioral measures, and provide recommendations on appropriate behavioral measures and metrics for future studies.
2. Improve methods for estimating dose-response functions for individual studies. This involves both developing and applying cutting-edge statistical methods, as well as considering what contextual variables in addition to acoustic dose can be incorporated into the analysis. The output will be improved estimates of response curves (with uncertainty) for each study.
3. Combine information across studies and species ("meta-analysis"), making use of expert biological opinion, to make predictions about taxa and contexts not yet studied. Differences in methods/protocols between studies will need to be accounted for. A component of this objective is to quantify the similarity/dissimilarity between species, placed into functional groups.
4. Based on the above, determine where major uncertainties still lie, and hence suggest where future experimental effort might be applied most fruitfully (through sensitivity analyses).

The first working group meeting took place in St Andrews between the 15$^{th}$ and 16$^{th}$ February 2012. This was the start-up meeting of the project and so involved all parties and included an overview of each BRS project, focusing particularly on the protocols used and metrics measured, and discussion of how to integrate data and results across projects. Participants in the February 2012 meeting are listed in Appendix 1.

## 3. Summary of first working group meeting

The first working group meeting started with a review of the main Behavioural Response Studies that have been conducted to date and included presentations about the availability of data from a range of sources, including opportunistic studies, as well as presentations on the main BRS projects: BRS Bahamas, BRS SOCAL, 3S and 3S2 and Cape Hatteras.

This allowed the group to then discuss the main methodological and taxonomic differences/similarities across studies. Whilst there has been considerable consistency in the technologies employed across all the BRSs, there have also been a range of important differences that need to be understood if data is to be collated and compared across studies. The most obvious difference is geographical location, resulting in differences in bathymetry, distance to coastline, sound propagation properties and prey availability (i.e., motivation to remain in area). In addition, the species that have been used in CEEs have varied both within and across studies. Differences also include sound types, sound frequencies, exposure durations, position/direction/movement of source relative to animal and distance to animal.

The group generated a list of contextual variables that need to be considered when combining data across taxon, trials and projects. A table was created based on this list to be populated for each exposure trial (Table 1, Appendix 2); this was identified as one of the first important tasks of the MOCHA project.

One of the characteristics of BRS data from each exposure is the diversity of metrics available, such as position, orientation, swim speed, dive time, social configuration and surfacing behavior. Within each BRS, methodologies have been applied to combine diverse behavioural metrics to characterise the behaviour of exposed animals and identify changes in it. Examples of analysis used include Principal Component Analysis approach used by the BRS SOCAL team or Mahanalobis distance metric used by the 3S team. The group discussed how to best combine these metrics, mainly in relation to the interpretation of output from the various metrics and the importance of understanding the broader consequences of responses identified.

Generally, participants agreed that it was important to relate behavioural changes to fitness consequences, and one possible avenue to investigate responses of other animals in the area would be to combine Dtag data with surface observations, particularly observations of conspecifics. However, the group recognised that the complexity in fitting models such as the bottlenose dolphin PCAD model was beyond the scope of MOCHA.

The 3S project presented the work it has been undertaking in developing dose-response curves. These rely on identified points of change in behaviour to estimate the dose at which animals responded. The results from 3S show that there can be a great deal of within-whale variation in response thresholds, when the same animal is exposed multiple times. A discussion followed where the group looked at different options to improve the model fitted to the dose-response curve that could be applied across all studies. The group considered integrating contextual variation, as this could explain some of the between and within-whale variation, associating dose with severity (incorporating expert opinion) rather than a binary response as used now, and parameterising models with timeline and severity data. Although dose-response curves do not relate behavioural changes to fitness consequences, the group felt that dose-response curves are still important outputs from these studies and needed by the community, particularly regulators. For example, they can be useful in describing the spatial extent of exposure and understanding the zone of influence. However, care should be taken with the definition of dose because received level could be interpreted very differently in a quiet versus noisy environment, suggesting that it would be important to take ambient noise measurements into account.

It was agreed that the MOCHA project should move forward simultaneously exploring the two approaches, by developing dose-response functions and looking at the change-point analysis (e.g., PCA and Mahalanobis). MOCHA Technical Groups were setup to develop these methodologies further (Appendix 3).

The group discussed methods to combine results across studies, particularly focusing on the aim to combine “raw” data to increase sample sizes and improve ability to extrapolate. In addition, methods developed by MOCHA should be able to help identify cases where there was no response, to provide confidence that the lack of response is true. Ground truthing of analytical methods developed by MOCHA could be achieved in the form of expert scoring, as developed by Southall and colleagues (2007) and the 3S project (Miller et al., in prep). Therefore a MOCHA Technical Group was set up to develop and carry out a consistent method of expert scoring all exposures from each project’s data set (Appendix 3). This could consist, for example, of each project setting up two independent scoring groups plus adjudicator and some cross-project validation of subsets of data. The expert scoring procedure should take into account contextual variation and how to quantify these variables. Another MOCHA Technical Group was setup that would discuss contextual variables, particularly those identified in Table 1 (Appendix 2).

Dtags are possibly the most consistent data acquisition technology used across various BRS studies and so may provide important covariates comparable across studies, although care should be taken when making comparisons across species because some parameters, such as flow noise, may vary due to different swimming speeds. There was also some discussion on levels of responses and how it is important to consider all possible responses, possibly using new methods that could deal with ordinal, multinomial data. A MOCHA Technical Group was set up to discuss state-space modelling and relating responses to life-history consequences (Appendix 3).

Finally the group discussed how to best proceed forward in combining methods across species. Focusing on species characteristics rather than species would allow generalisations to be made across functional groups. For example, dose-response models potentially could include several species, which could provide information on the relatedness of species’ responses without any prior categorisation of species into functional groups. In addition, it would be useful if models could provide information on the relative importance of context versus species in response levels. This approach could be tested within MOCHA during its initial 6 months using simulated data. Concern was raised that we might not have enough data to get informative species groupings resulting from the model, but the group agreed that it would be interesting to see if model selection reflects pre-conceived ideas about species groupings. In addition, individual projects agreed to make an effort to have cross-over in species should those opportunities to expose different species arise. The meeting ended with a discussion of the objectives for the next working group meeting.

## 4. Dates and objectives of next meeting

The second meeting of the working group will take place at the University of St. Andrews from 19th–21st September 2012. This meeting will focus on deep diver species (beaked whales and sperm whales[1]) because these include species of concern, there are data for these species across multiple BRS projects and the metrics measured are fairly well defined and social complexities are minimized. In preparation for this meeting, the MOCHA project will focus work on combining different available datasets and looking for common metrics that can be applied across species. A MOCHA Technical Group has been set up to prepare this meeting (Appendix 3).

**5. References**

Southall, B.L., Bowles, A.E., Ellison, W.T., Finneran, J.J., Gentry, R.L., Greene, C.R., Kastak, D., Ketten, D.R., Miller, J.H., Nachtigall, P.E., Richardson, W.J., Thomas, J.A., Tyack, P.L. (2007) Marine Mammal Noise Exposure Criteria: Initial Scientific Recommendations. Aquatic Mammals, 33(4).

Miller, P.J.O., Wensveen, P.J., Antunes, R., Alves, A.C., Lam, F.P.A., Visser, F., Tyack, P.L., Sivle, L.D., Kleivane, L., Kvadsheim, P. (in prep.) The severity of behavioral changes observed during experimental exposures of killer, long finned pilot, and sperm whales to naval sonar.

[1] Pilot whales are explicitly excluded because of their complex social structure, and are reserved for a future dedicated meeting

## APPENDIX 1 – First Working Group Meeting Participants

Len **Thomas**, CREEM, len@mcs.st-and.ac.uk
Catriona **Harris**, CREEM, catriona@mcs.st-and.ac.uk
Ian **Boyd**, SMRU, ilb@st-andrews.ac.uk
Patrick **Miller**, SMRU, pm29@st-andrews.ac.uk
Peter **Tyack**, SMRU, plt@st-andrews.ac.uk
Andy **Read**, DUML, aread@duke.edu
Petter **Kvadsheim**, FFI, Petter-Helgevold.Kvadsheim@ffi.no
Frans-Peter **Lam**, TNO, frans-peter.lam@tnl.nl
Brandon **Southall**, SEA, Brandon.Southall@sea-inc.net
Doug **Nowacek**, DUML, dpn3@duke.edu
John **Harwood**, CREEM/SMRU, jh17@st-andrews.ac.uk
Dave **Moretti**, NUWC, david.moretti@navy.mil
Ruth **King**, CREEM, ruth@mcs.st-and.ac.uk
Monique **MacKenzie**, CREEM, monique@mcs.st-and.ac.uk
Jason **Matthiopoulos**, CREEM, jm37@st-andrews.ac.uk
Jeremy **Goldbogen**, CRC, jgoldbogen@gmail.com
Stacy **DeRuiter**, CREEM, stacy_deruiter@yahoo.com
Dina **Sadykova**, CREEM, dinasadykova@live.com
Filipa **Samarra**, SMRU, fips2@st-andrews.ac.uk
Rene **Dekeling**, NLDMO, RPA.Dekeling@mindef.nl
Mike **Weise**, ONR, Michael.j.weise@navy.mil
Lise Doksaeter **Sivle**, IMR, lise.doksaeter.sivle@imr.no
Fleur **Visser**, KMR, fleurvisser@gmail.com
Ricardo **Antunes**, SMRU, rna@st-andrews.ac.uk
Paul **Wensveen**, SMRU, pwensveen@hotmail.com
Charlotte **Cure**, SMRU, charlotte.cure@laposte.net
Alex **Bocconcelli**, WHOI, abocconcelli@whoi.edu
Lars **Kleivane**, FFI, Lars.Kleivane@ffi.no
Gordon **Hastie**, SMRU, gdh@smru.co.uk

**Institutional addresses**


| | |
|---|---|
| CRC | Cascadia Research Collective, 218 1/2 W 4th Ave., Olympia, WA 98501, USA |
| CREEM | Centre for Research into Ecological and Environmental Modelling, University of St Andrews, St Andrews KY16 9LZ, UK. |
| DUML | Duke University Marine Lab, 135 Duke Marine Lab Rd Beaufort, NC 28516, USA |
| FFI | Maritime Systems Division, Norwegian Defence Research Establishment (FFI), NO-3191 Horten, Norway |
| IMR | Institute of Marine Research, Nordnesgaten 50, 5005 Bergen, Norway |
| KMR | Kelp Marine Research, Amsterdam, The Netherlands |
| NLDMO | NL Defence Materiel Organisation, Postbus 90822, 2509 LV Den Haag, Netherlands |
| NUWC | Navy Undersea Warfare Center, Newport, R.I. USA |
| ONR | Office of Naval Research, Marine Mammals & Biological Oceanography Program Code 322, One Liberty Center - Rm 1073, 875 N. Randolph St., Arlington, VA 22203-1995, USA |
| SEA | Southall Environmental Associates, Inc., 9099 Soquel Drive, Suite 8 Aptos, CA 95003, USA |
| SMRU | Sea Mammal Research Unit, Scottish Oceans Institute, University of St Andrews, St Andrews KY16 8LB, UK |
| TNO | Observation Systems, TNO, Defence, Security and Safety, The Hague, The Netherlands |
| WHOI | Woods Hole Oceanographic Institution, Woods Hole, MA 02543, USA |

**APPENDIX 2**

**Table 1. Contextual variables that differ across projects and exposure events**

| | | |
|---|---|---|
| **Focal animal** | Species | |
| | Age/sex class | |
| | History of exposure (exposure session – 1st or 2nd, and type of previous exposure – MF, Killer whale, PRN) | |
| | Habituation to/previous experience of noise (e.g. AUTEC versus 3S2 site) | |
| | Tagged or not (with one or multiple tags) | |
| **Source characteristics** | Signal type | |
| | Source type | |
| | Directionality | |
| | Depth of source | |
| | Frequency/Bandwidth | |
| | Source levels | |
| **Context** | Behavioural state at exposure point (and protocol) | |
| | Availability of prey/motivation data | |
| | Context in terms of land/bathymetry | |
| | Presence of other activities | |
| | Ambient noise levels | |
| | Presence of conspecifics | |
| **Exposure protocol** | Vessel moving? | |
| | Multiple playbacks on individual? | |
| | Timing and duration of ramp-up and exposure | |
| | Length of pre-exposure and post-exposure periods | |
| | Order of exposures | |
| | Timing of control period | |
| | Distance between source and animal | |
| | Received levels – measured? | |
| | Specificity of control sounds e.g. killer whales | |
| **Observation protocols** | Acoustic measurements/protocols | |
| | Visual observation protocols (tracking locations, group level behaviour) | |
| | Type of tag and measurements being collected by tag | |

## APPENDIX 3 – Initial MOCHA Technical Groups

- Response metrics and Change Point analysis
- Dose-response models applied to multi-species and including context covariates
- Expert scoring of behavioural responses
- Check exposure & response measurements and quantify contextual variables
- Relate responses to life history consequences (state-space modelling)
- Deep Diver Workshop Preparation